

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International BureauDoc Ref. AM5  
Appl. No. 09/458,299

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification<sup>6</sup> :</b> <b>A61K</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 96/22067</b> <b>(43) International Publication Date:</b> 25 July 1996 (25.07.96)
<b>(21) International Application Number:</b> PCT/US95/16290 <b>(22) International Filing Date:</b> 15 December 1995 (15.12.95)  <b>(30) Priority Data:</b> 08/366,332 27 December 1994 (27.12.94) US  <b>(71) Applicant:</b> UNITED BIOMEDICAL, INC. [US/US]; 25 Davids Drive, Hauppauge, NY 11788 (US).  <b>(72) Inventors:</b> KUEBLER, Peter, J.; P.O. Box 2114, Ft. Lauderdale, FL 33303-2114 (US). NIXON, Douglas, F.; 112 Bayview Avenue, Northport, NY 11768 (US).  <b>(74) Agent:</b> LIN, Maria, C., H.; Morgan & Finnegan, L.L.P., 345 Park Avenue, New York, NY 10154-0053 (US).		<b>(81) Designated States:</b> AU, CA, FI, JP, NO, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
<b>(54) Title:</b> PEPTIDE RATCHET LIBRARIES FOR CTL-INDUCING VACCINES AND THERAPEUTICS  <b>(57) Abstract</b>  The present invention relates to ratchet libraries composed of related peptides synthesized simultaneously in a single peptide synthesis. Ratchet libraries are derived from a longer template peptide by sequentially "ratcheting" the template sequence into the shorter ratchet length and are used for cytotoxic T lymphocyte (CTL) induction or stimulation if the CTL epitope is known. If the CTL epitope is unknown, then the ratchet library can be used for identification of CTL epitopes. The ratchet libraries can be prepared from any protein sequence to which an immune CTL response is desired and can be formulated for delivery as a vaccine or therapeutic for the treatment or prevention of disease or malignancy. For example, a ratchet library can be used in the prevention and treatment of infectious or malignant diseases including HIV, influenza, malaria, breast, ovarian, lung and colon cancers.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

-1-

1                   PEPTIDE RATCHET LIBRARIES FOR CTL-INDUCING  
2                   VACCINES AND THERAPEUTICS  
3  
4

5           FIELD OF THE INVENTION

6           The present invention relates to ratchet libraries  
7           composed of related peptides synthesized simultaneously in  
8           a single peptide synthesis. Ratchet libraries are derived  
9           from a longer template peptide by sequentially  
10          "ratcheting" the template sequence into the shorter  
11          ratchet length and are used for cytotoxic T lymphocyte  
12          (CTL) induction or stimulation if the CTL epitope is  
13          known. If the CTL epitope is unknown, then the ratchet  
14          library can be used for identification of CTL epitopes.  
15          The ratchet libraries can be prepared from any protein  
16          sequence to which an immune CTL response is desired and  
17          can be formulated for delivery as a vaccine or therapeutic  
18          for the treatment or prevention of disease or malignancy.  
19          For example, a ratchet library can be used in the  
20          prevention and treatment of infectious or malignant  
21          diseases including HIV, influenza, malaria, breast,  
22          ovarian, lung and colon cancers.

23          BACKGROUND OF THE INVENTION

24          The development of vaccines and therapeutics  
25          specifically designed to stimulate cytotoxic T lymphocytes  
26          (CTL) is needed. CTL are a vital component of the natural  
27          immune response against infectious organisms and malignant  
28          cells. CTL are CD8<sup>+</sup> thymus derived lymphocytes which  
29          appear early in an immune response and help in the  
30          elimination of, for example, virus-infected cells or tumor  
31          cells by lysis of the target cells and by secretion of  
32          chemical immunomodulators termed cytokines, such as  
33          interferons.

34          CTL have been detected following many viral  
35          infections, including HIV infection, and extensive  
36          evidence points to a major role for CTL in control of  
37          virus infections [McMichael et al. (1983) New Eng. J. Med.

-2-

1 301:13; Nixon et al. (1992) Immunology 76:515). For  
2 example, adoptive transfer of specific CTL to influenza  
3 [Taylor et al. (1986) Immunology 58:417] or paramyxovirus  
4 simian virus 5-infected mice, cleared the virus from the  
5 lungs [Young et al. (1990) J. Virol. 64:5403]. Tumor  
6 specific CTL have also been shown to clear tumors caused  
7 by mouse retroviruses [Cerundolo et al. (1987) Eur. J.  
8 Immunol. 17:173] and are also probably critical in the  
9 control of certain human malignancies. It has long been  
10 the aim of scientists to develop vaccines or therapeutics  
11 designed to specifically stimulate CTL immunity.

12 An essential step in the design of a CTL-inducing  
13 vaccine is in the identification of the antigenic sites to  
14 which CTL react. CTL recognize infected or malignant  
15 cells through the interaction of their specific T-cell  
16 receptor with a complex displayed on the surface of the  
17 target cell. The complex consists of an antigenic peptide  
18 specific to the virus or tumor, for example, and a major  
19 histocompatibility complex (MHC) class I molecule encoded  
20 by the Class I MHC genes of the host [Townsend et al.  
21 (1986) Cell 44:959]. Clusters of closely linked MHC  
22 alleles are characteristically inherited as a genetic unit  
23 termed the "haplotype". In general, individual MHC  
24 molecules associate with and present different antigenic  
25 protein fragments, so that one fragment of an antigenic  
26 protein is recognized by CTL of a specific MHC haplotype,  
27 while a different MHC haplotype requires another fragment  
28 of the antigen for recognition, i.e., recognition of  
29 individual antigenic fragments is MHC-restricted. As the  
30 MHC alleles are highly polymorphic between diverse genetic  
31 groups, a large number of distinct peptides may be needed  
32 to insure CTL stimulation across diverse human  
33 populations.

34 The exact fragment(s) of a virus or tumor antigen or  
35 other potential antigenic site (i.e., CTL epitope)  
36 recognized by a specific CTL was thought to be between 7

-3-

1 and 25 amino acids, but recent characterization of viral  
2 peptides naturally processed in virus-infected cells and  
3 displayed by Class I MHC molecules have identified the CTL  
4 epitopes as peptides of between 7 to 11 amino acids in  
5 length [Rötzschke et al. (1991) Immunology Today 12:447]  
6 with the majority of these peptides being of 9 amino acids  
7 (nonomers).

8 Identification of CTL epitopes in a protein sequence  
9 has been achieved by using synthetic peptides to map  
10 immunogenic sites. For example, several human CTL epitopes  
11 have been defined from HIV through an *in vitro* testing  
12 process of the human immune response to HIV infection  
13 [Nixon et al. (1988) Nature 336:484-487; Nixon et al. U.K.  
14 Patents GB 2,255,093, 2,273,709, 2,273,710]. While many  
15 HIV CTL epitopes have been identified in animals, few have  
16 been identified in humans. However, because CTL epitopes  
17 are simultaneously recognized by a T-cell receptor that is  
18 specific for both the virally-encoded peptide and the  
19 host-encoded MHC for clearance of an infected cell to  
20 occur, CTL-epitopes are species specific. Hence, human  
21 CTL epitopes may not be reliably predicted from animal  
22 studies.

23 While vaccine development has led to successful  
24 vaccine against many infectious diseases, (e.g. polio,  
25 measles), there are several important pathogens for which  
26 vaccines are either ineffective or simply non-existent,  
27 for example HIV, hepatitis C virus (HCV) and herpes  
28 simplex virus (HSV). Moreover, there are no vaccines for  
29 treatment of malignancies.

30 The identification of CTL epitopes makes it feasible  
31 to design CTL-stimulating vaccines and other  
32 immunotherapeutics for prevention or treatment of disease  
33 by the clearance of virally-infected cells or malignant  
34 cancer cells. However, there remain at least four major  
35 problems associated with developing CTL-inducing vaccines  
36 and immunotherapeutics, namely, (1) the identification of

-4-

1 CTL epitopes or regions of proteins containing such  
2 epitopes, (2) the induction of specific CTL responses by  
3 peptides, (3) the need to accomodate MHC diversity with a  
4 large multiplicity of peptides and (4) the ability to  
5 provide for antigenic variation and escape mutations  
6 within the CTL epitopic regions.

7 With respect to the identification of CTL epitopes,  
8 progress has been made in identifying CTL epitopes from a  
9 number of target antigenic proteins,; however, there  
10 remain many proteins which contain potential CTL antigenic  
11 sites for which epitopes have not been identified. For  
12 example, the EBNA 1 protein of Epstein-Barr virus (EBV).  
13 The present invention provides a solution to this problem  
14 because the ratchet libraries can encompass extensive CTL  
15 antigenic regions and eliminate the need to precisely map  
16 CTL epitopes, or even to map the CTL epitopes at all. In  
17 addition, the ratchet libraries can be used to map  
18 antigenic sites.

19 Until recently, it was assumed that the induction of  
20 specific CTL responses by peptides could only be  
21 stimulated by endogenously-produced peptide fragments of  
22 endogenous proteins assembled into HLA Class I-antigenic  
23 peptide complexes on the cell surface. However, recent  
24 studies have demonstrated that CTL responses can be primed  
25 by administration of lipid-derivatized peptides [Deres et  
26 al. (1989) Nature 342:561], peptides in liposomes [Friede  
27 et al. (1994) Vaccine 12:791-797], or peptides admixed or  
28 conjugated to other biologically active substances [Shirai  
29 et al. (1994) J. Immunol. 152:549]. Hence, ratchet  
30 library peptides can be formulated into an appropriate  
31 vehicle to elicit CTL responses in vivo.

32 To accomodate genetic diversity and MHC restriction,  
33 the ratchet library peptides provide a major advance since  
34 several epitopes can be incorporated into the ratchet  
35 libraries rather than relying upon mixtures individually  
36 synthesized immunogenic peptides.

-5-

1           The ability to provide for antigenic variation and  
2           escape mutations within the CTL epitopic regions is  
3           another significant problem. Antigenic variation is a  
4           recurrent problem among certain pathogens contributing to  
5           unsuccessful or limited success of vaccines. Extensive  
6           antigenic variation, for example, is a hallmark of HIV  
7           (AIDS), rhinovirus (the common cold), influenza virus  
8           (flu), plasmodium falciparum (malaria). In addition, some  
9           tumors and infectious agents use utilized escape mutation  
10          to avoid immune surveillance. Ratchet libraries can be  
11          constructed to embody known antigenic variation and escape  
12          mutations to pre-empt these problems.

13          Hence, the ratchet library method of CTL induction  
14          provides a solution to obstacles in the development of  
15          vaccines and therapeutics for such pathogens or cancers.

#### 16 17          SUMMARY OF THE INVENTION

18          This invention is directed to a ratchet library of  
19          peptides comprising at least one immunostimulatory  
20          cytotoxic T lymphocyte (CTL) epitope. The peptides are of  
21          length  $l$ . The sequences of the peptides in the library  
22          are determined from a template peptide of length from  $l+1$   
23          to  $n$  amino acids such that each position  $x$  in the library  
24          has all the amino acids present in the template peptide at  
25          positions  $x$  to  $n-(l-x)$ , inclusive and the ratio of amino  
26          acids at each position  $x$  is determined by the relative  
27          prevalence of amino acids at that position  $x$ . In  
28          accordance herewith  $l$  is from about 7 to about 25 amino  
29          acids, preferably 8-10 and more preferably 9;  $n$  is from  
30           $l+1$  to about 100, preferably from  $l+1$  to about 75 and more  
31          preferably from  $l+1$  to about 50; and  $x$  is from 1 to  $l$ .

32          Moreover, if a position  $x$  is identified as part of an  
33          MHC-binding motif of a CTL epitope, then that position  $x$   
34          in the ratchet library is fixed as one or more amino acids  
35          of the MHC-binding motif in an equimolar ratio.

-6-

1           The CTL epitopes of this invention are from a virus,  
2           bacterium, parasite, tumor antigen, allergen or other  
3           protein antigen. The ratchet library can be constructed  
4           from the template peptides of any one of SEQ ID NOS: 1-11.  
5           the peptides can have a covalently attached N-terminal  
6           tripalmitoyl-5-glycerol-cysteine moiety or be linked to a  
7           branched core sequence, polymerized or conjugated to a  
8           carrier molecule.

9           Another aspect of this invention provides a  
10          pharmaceutical or vaccine composition comprising the  
11          subject ratchet libraries including emulsion or a  
12          microparticle formulation with or without the addition of  
13          free Pam<sub>3</sub>Cys or a derivative thereof. These compositions  
14          are useful in a method of treating or preventing a disease  
15          or a malignancy which comprises administering an amount of  
16          the ratchet library as a vaccine or pharmaceutical  
17          composition to a mammal effective to stimulate a CTL  
18          response against the disease or the malignancy associated  
19          with the CTL epitope present in the library.

20          Still another aspect of the invention is directed to  
21          a method of constructing a library of related peptides to  
22          provide a ratchet library which comprises identifying a  
23          template peptide; calculating a distribution of amino  
24          acids at each position  $x$  having those amino acids present  
25          in the template peptide at positions  $x$  to  $n-(1-x)$ ,  
26          inclusive, wherein  $1$  is from about 7 to about 25,  $n$  is  
27          from  $1+1$  to about 100, and  $x$  is from 1 to 1; and  
28          synthesizing said ratchet library.

#### 29           BRIEF DESCRIPTION OF THE DRAWINGS

30           Fig. 1 depicts malaria ratchet libraries 1 and 2 from  
31           *Plasmodium berghei* circumsporozoite (CS) protein and their  
32           construction. Fig 1A shows the template peptide with the  
33           known CTL epitope (CS 252-260) indicated by a box. Below  
34           the template peptide is the corresponding set of  
35           overlapping nonmer peptides. Fig. 1B provides an example  
36

-7-

1 of a sequence alignment for malaria ratchet library 1 (top  
2 panel), its corresponding amino acid (AA) distribution by  
3 position (middle panel) and the percent of each amino acid  
4 at each position in the library (bottom panel).

5 Fig. 2 is a graphic illustration of malarial-specific  
6 CTL activity against CS 252-260. The graph shows the  
7 percent specific cell lysis as a function of the effector  
8 to target cell (E:T) ratio in mice immunized with 100  $\mu$ g  
9 doses of malaria ratchet library 1 in microparticles.

10 Fig. 3 is a graphic illustration of malarial-specific  
11 CTL activity against CS 252-260. The graph shows the  
12 percent specific cell lysis as a function of E:T ratio in  
13 mice immunized with 1 mg doses of malaria ratchet library  
14 2 in microparticles.

15 Fig. 4 is a graphic illustration of malarial-specific  
16 CTL activity against CS 247-266. The graph shows the  
17 percent specific cell lysis as a function of E:T ratio in  
18 mice immunized with 1 mg doses of malaria ratchet library  
19 2 in microparticles.

20 Fig. 5 is a graphic illustration of malarial-specific  
21 CTL activity against CS 252-260. The graph shows the  
22 percent specific cell lysis as a function of E:T ratio in  
23 mice immunized with 10  $\mu$ g doses of malaria ratchet library  
24 1 as lipopeptides.

25 Fig. 6 is a graphic illustration of the lack of  
26 malarial-specific CTL activity against a self peptide.  
27 The graph shows the percent specific cell lysis as a  
28 function of E:T ratio in mice immunized with 100  $\mu$ g doses  
29 of malaria ratchet library 1 as lipopeptides.

30 Fig. 7 depicts an MHC-restricted malaria ratchet  
31 library constructed from malarial ratchet library 1 (top  
32 panel). The middle panel shows the known anchor residues  
33 for four MHC haplotypes, K<sup>d</sup>, D<sup>b</sup>, K<sup>b</sup> and L<sup>d</sup>. These anchor  
34 residues are at positions 2 for K<sup>d</sup> and L<sup>d</sup>, 5 for D<sup>b</sup>, K<sup>b</sup> and  
35 L<sup>d</sup>, 8 for K<sup>b</sup> and 9 for K<sup>d</sup>, D<sup>b</sup> and L<sup>d</sup>. The bottom panel

-8-

1 shows the resulting MHC-restricted malaria ratchet  
2 library.

3 Fig. 8 depicts an HIV ratchet library from a 35 amino  
4 acid sequence of the HIV-1 gp120 V3 loop region (residues  
5 305-339). Fig. 8A provides the sequence of 15 HIV-1  
6 variants from this region (top panel) and the  
7 corresponding SSAL from those sequences. The D<sup>d</sup>  
8 restricted CTL epitope at gp120 amino acid positions 318-  
9 326 is indicated by a box. Fig. 8B provides the amino  
10 acid (AA) distribution by position (middle panel) and the  
11 percent of each amino acid at each position in the library  
12 (bottom panel) in a ratchet library constructed from the  
13 SSAL. Amino acids divergent from the consensus B sequence  
14 are shown as upper case letters and conserved amino acids  
15 in the consensus sequences are shown as lower case  
16 letters.

17 Fig. 9 depicts an HIV-1 gag peptide linear ratchet  
18 library containing a mouse HIV CTL epitope prepared from a  
19 100 amino acid template (top panel). The amino acid (AA)  
20 distribution by position (middle panel) and the percent of  
21 each amino acid at each position in the library (bottom  
22 panel) is shown. The D<sup>b</sup> restricted epitope (at gag  
23 residues 390-398) is indicated by the box.

24 Fig. 10 depicts an HIV-1 gag peptide linear ratchet  
25 library containing a mouse HIV CTL epitope prepared from a  
26 40 amino acid template (top panel). The amino acid (AA)  
27 distribution by position (middle panel) and the percent of  
28 each amino acid at each position in the library (bottom  
29 panel) is shown. The D<sup>b</sup> restricted epitope (at gag  
30 residues 390-398) is indicated by the box.

31 Fig. 11 depicts an HIV-1 gag peptide linear ratchet  
32 library containing a mouse HIV CTL epitope prepared from a  
33 20 amino acid template (top panel). The amino acid (AA)  
34 distribution by position (middle panel) and the percent of  
35 each amino acid at each position in the library (bottom

-9-

1 panel) is shown. The D<sup>b</sup> restricted epitope (at gag  
2 residues 390-398) is indicated by the box.

3 Fig. 12 is a graphic illustration of HIV-specific CTL  
4 activity against HIV-1 gag residues 390-398. The graph  
5 shows the percent specific cell lysis as a function of E:T  
6 ratio in mice immunized with 100 µg doses of the HIV  
7 ratchet library from the 40-mer template in an emulsion.

8 Fig. 13 depicts a mucin ratchet library from a 20-mer  
9 repeating sequence (top line). The next two lines  
10 illustrate two alternate template peptides for  
11 construction of this mucin ratchet library. The boxed  
12 residues are the additional sequences added at the termini  
13 to allow representation of all possible nonomers of the 20  
14 amino acid repeat sequence. The amino acid (AA)  
15 distribution by position (middle panel) and the percent of  
16 each amino acid at each position in the library (bottom  
17 panel) is shown.

18 Fig. 14 depicts a mutant p53 ratchet library  
19 constructed from a template peptide of amino acids 124-151  
20 (top panel). There are additional amino acids, which  
21 represent known p53 mutants incorporated at positions 9-  
22 12. The boxed residues are 10 amino acid CTL epitope  
23 identified in Balb/C mice. The amino acid (AA)  
24 distribution by position (middle panel) and the percent of  
25 each amino acid at each position in the library (bottom  
26 panel) is shown.

27 Fig. 15 depicts influenza ratchet library 1  
28 constructed from a 25-mer template sequence of residues  
29 139-163 of influenza A A/34/PR8 nucleoprotein (top panel).  
30 The known K<sup>d</sup>-restricted epitope of residues 147-155 is  
31 indicated by a box. The amino acid (AA) distribution by  
32 position (middle panel) and the percent of each amino acid  
33 at each position in the library (bottom panel) is shown.

34 Fig. 16 depicts influenza ratchet library 2 from a  
35 template peptide which is a linkage of 3 CTL epitopes in  
36 order from N to C terminus of residues 50-58, 147-155 and

-10-

1 366-374 of the nucleoprotein. Each CTL epitope is  
2 indicated by a box. The amino acid (AA) distribution by  
3 position (middle panel) and the percent of each amino acid  
4 at each position in the library (bottom panel) is shown.  
5

#### 6 DETAILED DESCRIPTION OF THE INVENTION

7 The present invention is directed to a ratchet  
8 library of peptides comprising at least one  
9 immunostimulatory cytotoxic T lymphocyte (CTL) epitope.  
10 The peptides of the ratchet library are of length  $l$  and  
11 the sequences of the peptides in the library are  
12 determined from a template peptide having a length from  
13  $l+1$  to  $n$  amino acids. Each position  $x$  in the library  
14 peptides has those amino acids which are present in the  
15 template peptide at positions  $x$  to  $n-(l-x)$ , inclusive.  
16 Accordingly, the ratio of the individual amino acids at  
17 each position  $x$  is determined from the relative numbers  
18 (or prevalence) of the amino acids at that position  $x$ . In  
19 accordance with this invention,  $l$  is from about 7 to about  
20 25,  $n$  is from  $l+1$  to about 100, and  $x$  is from 1 to  $l$ .

21 For example, in a ratchet library of the above  
22 formula, position 1 contains all the amino acids of the  
23 template peptide at positions 1 to  $n-(x-1)$ , position 2  
24 contains all the amino acids of the template peptide at  
25 positions 2 to  $n-(x-2)$ , position 3 contains all the amino  
26 acids of the template peptide at positions 3 to  $n-(x-3)$ ,  
27 ... , position  $x-1$  contains all the amino acids of the  
28 template peptide at positions  $x-1$  to  $n-1$ , and position  $x$   
29 contains all the amino acids of the template peptide at  
30 positions  $x$  to  $n$ .

31 Fig. 1 provides an example of a malaria ratchet  
32 library, showing the relative ratio of amino acids at each  
33 position in the ratchet library derived from a longer  
34 malaria template peptide as well as the percentage of  
35 amino acids required for synthesis at each position of the  
36 ratchet library. More specifically, the malaria template

-11-

1 peptide is divided up into sequential 9-mers which are  
2 aligned at the amino terminus of the template peptide to  
3 create a set of overlapping peptides from which to  
4 calculate the ratchet library composition. After  
5 calculation of its composition, the ratchet library is  
6 prepared in a single synthesis based on the calculated  
7 amino acid distributions at each position.

8 The size of the ratchet l, or ratchet length, can be  
9 determined from the actual or approximate size of the  
10 target CTL epitope. CTL epitopes have been identified  
11 which vary in length from 7 to 25 residues. However, the  
12 majority of CTL peptides are from 8-10 amino acids, and  
13 many are 9 amino acids. While the actual size of the CTL  
14 inducing peptide is preferred to determine the length l of  
15 the ratchet library, e.g. 9 amino acids, the ratchet  
16 length can be determined by other means, including an  
17 arbitrary selection of size within the range of 7 to 25  
18 amino acids.

19 The size of the template peptide ranges from l+1 to  
20 about 100 amino acids, and preferably from l+1 to about 75  
21 amino acids, and more preferably from l+1 to about 50.  
22 Selection of the template peptide length is an important  
23 factor in determining the overall complexity of the  
24 ratchet library, so that shorter template peptides tend to  
25 yield less complex ratchet libraries, i.e., have fewer  
26 peptides in the library.

27 If the CTL epitope is known, then the template  
28 peptide should be of a length n such that the CTL epitope  
29 is flanked by sufficient adjoining sequences, preferably  
30 at least l-1, to insure that the CTL epitope is  
31 represented in the ratchet library. If the CTL epitope is  
32 not known, then the template peptide can have a length  
33 which covers a significant region of the protein being  
34 tested. Typically such a length can range from about 20  
35 to about 100 residues, and preferably ranges up to 50 or  
36 75 residues. The template peptide can also be selected on

-12-

1 the basis of clustering of epitopes, of hydrophobicity, of  
2 stretches containing basic amino acids or of another  
3 biological characteristic. Selection of a longer template  
4 peptide is useful in identifying unknown CTL epitopes.

5 MHC binding motifs have identified particular amino  
6 acids residues within CTL epitopes which are important in  
7 peptide binding to the MHC receptor. When the MHC binding  
8 motif is known for a particular CTL epitope, then the  
9 ratchet library can be simplified by replacing the  
10 calculated distribution of amino acids at a particular  
11 site with the ratio of known amino acids from the MHC  
12 binding motif at that site. An example of this is shown  
13 in Example 2. In another example, the proportion of amino  
14 acids within a ratchet library can be altered to reflect  
15 human HLA binding motifs. For example, human HLA binding  
16 motifs for 9-mer or 10-mer peptides typically have the  
17 designated amino acids at the indicated positions: for  
18 HLA-A2, leucine at position 2, valine or leucine at the C  
19 terminus; for HLA-B35, proline at position 2, tyrosine at  
20 the C terminus; for HLA-B53, proline at position 2,  
21 phenylalanine or tryptophan at the C terminus; for HLA-B8,  
22 lysine at position 3, lysine at position 5, isoleucine at  
23 the C terminus; for HLA-B27, arginine at position 2,  
24 lysine or arginine at the C terminus; for HLA-B7, alanine  
25 at position 1, proline at position 2, arginine at position  
26 3, leucine or valine at the C terminus; for HLA-A68,  
27 threonine or valine at position 2, arginine at the C  
28 terminus; for HLA-A3.1, isoleucine or leucine at position  
29 2, phenylalanine at position 3, lysine or tyrosine at the  
30 C terminus; and for HLA-A11, isoleucine or leucine at  
31 position 2, lysine at the C terminus.

32 When the CTL epitope is contained in a multiple  
33 tandemly repeated sequence, then the template peptide  
34 length n can be equal to the length of the CTL epitope  
35 plus 1-1 residues where the 1-1 residues carboxy-terminal  
36 amino acids of the epitope are placed at its amino

-13-

1 terminus, or alternatively, the 1-1 amino-terminal  
2 residues of the epitope are placed at its carboxyl  
3 terminus, for example as shown in Fig. 5 for the mucin  
4 ratchet library.

5 To accomodate antigenic variation, the ratchet  
6 library can be constructed from a template peptide which  
7 is itself a "structured synthetic antigen library" or  
8 SSAL. SSALs are defined and exemplified in U.S. Serial  
9 No. 08/143,412, filed October 26, 1993, which is  
10 incorporated herein by reference. Briefly, the sequence  
11 of an SSAL is determined by aligning the primary amino  
12 acid sequences of a related family of CTL epitopes and  
13 identifying the invariant and variant loci within the  
14 alignment. The invariant loci generally represent the  
15 structural framework of the SSAL. The degeneracy within  
16 the SSAL is determined by the loci within the alignment  
17 that harbor different amino acid residue types relative to  
18 an arbitrary prototype sequence. After determining which  
19 amino acids are to be at each position, the degree of  
20 degeneracy for the multiresidue position in the SSAL  
21 library is determined from the number of variants each  
22 individual amino acid represents by one of three methods.  
23 Thus in a simple manner, the specific amino acids and  
24 their frequency of appearance at each position within the  
25 SSAL is defined by the primary sequences of the different  
26 CTL antigens or molecules in the alignment of multiple  
27 primary sequences.

28 The degeneracies for the variant amino acid positions  
29 used for an SSAL can be determined in one of three ways.  
30 In one method, the identity and ratio of residues is  
31 determined by the relative prevalence of the amino acids  
32 in a compilation of known sequences for the epitope. In  
33 another method the identity of the amino acids at the  
34 variant position is determined from the compilation of  
35 known sequences for the epitope but the ratio of amino  
36 acids is set to be equimolar. Finally, the identity and

-14-

1 ratio of amino acids at a variant position can be  
2 determined by a modification of the first method to  
3 provide a simplified SSAL. When some of the residue types  
4 are present at a low frequency (i.e., less than 5-10%  
5 representation), the complex SSALs are modified to ensure  
6 adequate representation of all variants. The process  
7 follows three rules: 1) any amino acid present in the  
8 primary sequence list at a proportion less than 10% is set  
9 at 5% to allow for adequate representation of all variable  
10 positions; 2) amino acids occurring at frequencies greater  
11 than or equal to 10% are rounded to the nearest 10% of  
12 prevalence. If the sum of percent prevalence exceeds  
13 100%, the percent of the amino acid with the highest  
14 prevalence is correspondingly reduced so that the amino  
15 acids occurring at a given position are each represented in  
16 the SSAL but the total representation does not exceed  
17 100%.

18 Once an SSAL is calculated, then it can be the  
19 "template peptide" for construction of a ratchet library  
20 in accordance with the formula such that each position x  
21 in the library peptides has those amino acids which are  
22 present in the template SSAL at positions x to n-(1-x),  
23 inclusive. In other words, each full composition and  
24 ratios of amino acids at each position x in the SSAL is  
25 "ratcheted" and used to calculate the final distribution  
26 of amino acids in the ratchet library.

27 The ratchet libraries can be prepared as a ensemble  
28 of linear peptides. Similarly, it can be attached to a  
29 branched core sequences, conjugated to a carrier or  
30 polymerized.

31 These core sequences include dendritically branched  
32 cores, linear array type branched cores or randomly  
33 branched cores (e.g. poly-L-lysine). The branched cores  
34 can be composed of an amino acid or an amino acid analog  
35 having two amino groups and one carboxyl group, each group  
36 capable of forming a peptide bond linkage. Preferably

-15-

1 such amino acids are lysine or a lysine analog such as  
2 ornithine. The amino acid analog can be an  $\alpha$ -amino acid,  
3 a  $\beta$ -amino acid, or any other either natural or non-natural  
4 amino acid with two amino groups and one carboxyl group  
5 available for forming peptide bonds. Preferred branched  
6 peptides of the invention are dimers, tetramers and  
7 octamers, especially those having a branching core  
8 structure composed of lysine such as a heptalysine core.  
9 Similarly, the branched cores can contain other residues  
10 interspersed among the branching residues as depicted, for  
11 example, in Fig. 12 of U.S. Serial No. 143,412.

12 When branched ratchet libraries are made, the library  
13 can have a C-terminal methionine as the residue that is  
14 attached to the branched core. The methionine provides a  
15 cleavable site to facilitate analysis of the ratchet  
16 library.

17 In addition, the ratchet can have one or more lysine  
18 residues (added at the amino terminus) to increase peptide  
19 solubility, cysteine and haloacylated residues can be  
20 added to facilitate directed coupling to carrier  
21 molecules, and methionine can be added for cyanogen  
22 bromide cleavage if necessary. Pam,Cys, or a similar  
23 lipid tail, can be added to create a lipopeptide.

24 The subject ratchet libraries can also be used to  
25 form conjugates, i.e., the ratchet library, either in  
26 branched or linear form, can be coupled directly or  
27 indirectly, by methods known in the art, to carriers such  
28 as bovine serum albumin (BSA), human serum albumin (HSA),  
29 or other proteins, red blood cells or latex particles. In  
30 another embodiment, a ratchet library can be polymerized  
31 to homo- or hetero-dimers or higher oligomers by cysteine  
32 oxidation, by induced disulfide cross-linking, or by use  
33 of homo- or hetero-functional multivalent cross-linking  
34 reagents.

35 As used herein, a CTL epitope is a fragment of an  
36 antigen which binds to the peptide-binding cleft of an MHC

-16-

1 molecule such that the fragment-MHC complex is recognized  
2 by a T cell antigen-specific receptor (TCR) and thereby  
3 stimulates a CTL response. CTL and T cell epitopes are  
4 reviewed in Encyclopedia of Immunology (Roitt et al.,  
5 eds.), 1992, Academic Press Ltd, London, in Vol. I at  
6 pages 447-450 and pages 515-517, respectively.

7 The protein sequence selected for a ratchet library  
8 can be from a protein with known immunogenic CTL epitopes,  
9 or from a protein whose CTL-stimulating ability has not  
10 been determined, in which case the ratchet library method  
11 can be used to identify CTL epitopes. Ratchet libraries  
12 can be constructed from CTL epitopes (or putative CTL  
13 epitopes ) of viruses, bacteria, parasites, tumor  
14 antigens, allergens, amino acid sequences deduced from an  
15 intron or exon/intron mixtures, or from aberrant proteins  
16 often associated with malignancy and generated by  
17 frameshift mutations (i.e. frameshift sequences), or any  
18 other proteins known to stimulate a CTL response. More  
19 specifically, ratchet libraries can be prepared from the  
20 following proteins or proteins from the listed organisms  
21 or diseases (with the cited references indicating known  
22 CTL response to those proteins): melanoma proteins  
23 [Bakker et al. (1994) J. Exp. Med. 179:1005] including  
24 MAGE-1, -2, and -3 [Gaugler et al. (1994) J. Exp. Med.  
25 179:921]; proteins associated with renal cell carcinoma;  
26 proteins associated with colon carcinoma [Townsend et al.  
27 (1994) Nature: 371:662]; proteins associated with prostate  
28 cancer (malignant or benign) including PSA; tyrosinase  
29 [Brichard et al. (1993) J. Exp. med. 178:489]; oncogenes  
30 such as the HER-2/neu proto-oncogene; ras [Gedde-Dahl et  
31 al. (1994) Eur. J. Immunol. 24:410]; MUC1 [Barnd et al.  
32 (1989) Proc. Natl. Acad. Sci. USA 86:7159]; p53 [Mijman et  
33 al. (1994) Immunol. Lett. 40:171]; p16; TL [Morita et al.  
34 (1994) J. Exp. Med. 179:777]; proteins from HIV-1 or -2  
35 including envelope, gag, pol, nef, tat, rev, vpx, vpu  
36 [Nixon et al. (1992)]; HTLV-I or -II including envelope,

-17-

gag, pol, pX and TAX [Jacobson et al. (1990) Nature 348:245; lymphocytic choriomeningitis virus of mice (LCMV) [Aebischer et al. (1991) Proc. Natl. Acad. Sci. USA 88:11047]; influenza A, B or C including PB1, PB2, PA, NS1, M1, NP, HA [McMichael et al. (1978) Eur. J. Immunol. 8:705]; Epstein-Barr virus (EBV) including TETA, EENL, EBNA3, EBNA1 and LMP [Brooks et al. (1993) J. Exp. Med. 178:879]; respiratory syncytia virus (RSV) [Bangham et al. (1985) J. Virol. 56:55]; hepatitis B virus (HBV) [Bertoletti et al. (1993) J. Virol. 67:2367]; hepatitis C virus (HCV) [Koziel et al. (1992) J. Immunol. 149:3339]; herpes simplex virus (HSV) [Bonneau (1993) Virology 195:62]; cytomegalovirus (CMV) [Borysiewicz et al. (1988) Eur. J. Immunol. 18:269]; parainfluenza virus 1 including hemagglutinin, neuraminidase, phosphoprotein and nucleoprotein [Dave et al. (1994) Virology 199:376]; intracisternal A particle gag [de Bergeyck et al. (1994) Eur. J. Immunol. 24:2203]; bovine leukemia virus [Gatei et al. (1993) J. Virol. 67:1796]; papilloma viruses [Feltkamp et al. (1993) Eur. J. Immunol. 23:2242]; malaria including *P. falciparum*, *P. berghei*, *P. ovale*, *P. vivax*, *P. malaria* [Aggarwal et al. (1990) J. Exp. Med. 172:1083]; *Histoplasma capsulatum* [Deepe (1994) J. Immunol. 152:3491]; *Listeria* [Harty et al. (1992) J. Exp. Med. 175:1531]; Toxoplasmosis [Khan et al. (1994) J. Immunol. 152:1856]; *Trypanosoma cruzi*; *Yersinia*; *M. tuberculi*; *M. lepri*; *Pneumocystis carinii*; Kaposi's sarcoma or frameshift sequences [Townsend (1994)].

The preferred ratchet libraries of this invention are those libraries provided in the Examples and the Figures.

CTL responses can be measured by conventional techniques known to ordinarily skilled artisans, including, for example, the dye exclusion test and the Cr-release assay described in Encyclopedia of Immunology, supra at page 451. Another method to assay CTL is

-18-

1 described by McDonald et al. (1980) Immunol. Rev. 51:93-  
2 123.

3 The ratchet libraries are prepared by chemical  
4 synthesis using standard techniques well known in the art  
5 such as the solid-phase synthetic route pioneered by  
6 Merrifield. The coupling of multiple amino acids at a  
7 given position is accomplished by providing a mixture of  
8 the desired amino acids at the ratios determined by the  
9 ratchet process. If necessary the ratio of amino acids in  
10 the mixture can be varied to account for different  
11 coupling efficiency of those amino acids.

12 Based on CTL induction of the ratchet libraries, they  
13 are useful in a vaccine composition to treat or prevent  
14 disease or malignancy in accordance with the source of the  
15 CTL epitope in the ratchet library. In other words an HIV  
16 ratchet library can be used as an HIV CTL vaccine (either  
17 as a vaccine component or as a therapeutic in the  
18 treatment of AIDS), an HCV ratchet library as an HCV CTL  
19 vaccine, an influenza ratchet library as a flu CTL  
20 vaccine, a mutant p53 ratchet library as a cancer CTL  
21 vaccine and the like.

22 For example, efforts to develop a malaria vaccine  
23 have been hampered by the complexity of the parasite life  
24 cycle and the inability for an antibody-inducing vaccine  
25 alone to provide sufficient efficacy. A CTL response to  
26 liver stage antigens of the malaria parasite *Plasmodium*  
27 *falciparum* has been recently reported [Hill et al. (1992)  
28 Nature 360:434]. Stimulation of the CTL response against  
29 the parasite appears necessary for an effective vaccine,  
30 since CTL can eliminate parasite-infected cells at the  
31 liver stage when the parasite load is low.

32 Vaccine compositions containing one or more distinct  
33 ratchet libraries can be introduced into normal subjects  
34 to stimulate production of CTL by immunization protocols  
35 known in the art. Similarly the subject ratchet libraries  
36 (one or more libraries) can be formulated in a vaccine

-19-

1 composition using adjuvants, pharmaceutically-acceptable  
2 carriers or other ingredients routinely provided in  
3 vaccine compositions. Adjuvants for use in this invention  
4 include incomplete Freund's adjuvant (IFA), alum, lipidic  
5 amino acids and Pam<sub>3</sub>Cys (see description in the Examples).  
6 These latter two adjuvants can be either covalently  
7 attached to the ratchet to produce a lipopeptide ratchet  
8 library or formulated together with the ratchet library  
9 for co-administration.

10 Vaccine formulations are readily determined by one of  
11 ordinary skill in the art and include formulations for  
12 immediate release and for sustained release. Formulations  
13 contemplated by this invention include microparticles,  
14 microcapsules, emulsions, liposomes, DMSO-glycerol and the  
15 like.

16 The present vaccines can be administered by any  
17 convenient route including subcutaneous, oral,  
18 intramuscular, intravenous, intra-dermal, intraocular,  
19 vaginal, trans-dermal or other parenteral or enteral  
20 route. Similarly the vaccines can be administered as a  
21 single dose or divided into multiple doses for  
22 administration.

23 The vaccine compositions of the instant invention  
24 contain an immunoeffective amount of a ratchet library to  
25 treat or prevent disease or malignancy associated with the  
26 source of the CTL epitope in that ratchet library.  
27 Preferred vaccine compositions are effective for CTL  
28 induction with respect to malaria, HIV, HCV, mucin, p53  
29 and influenza and their associated pathogenic conditions.  
30 Such compositions in dosage unit form can contain about 10  
31 ng to about 2 mg of the peptide (or mixture of peptides)  
32 per kg body weight. When delivered in multiple doses, the  
33 dosage unit form is conveniently divided into the  
34 appropriate amounts per dosage.

35 Accordingly, another aspect of this invention  
36 provides a method of treating or preventing a disease or a

-20-

1 malignancy which comprises administering an amount of the  
2 library of Claim 1 to a mammal effective to stimulate a  
3 CTL response against the disease or malignancy associated  
4 with a CTL epitope present in said library. Based on the  
5 source of the ratchet library (i.e., the virus, bacterium  
6 or other organism from which the template peptide was  
7 derived or a protein associated with malignancy), then one  
8 skilled in the art can readily determine the amount needed  
9 for delivery to obtain the desired therapeutic result,  
10 that is, the amount of library to induce a CTL response of  
11 therapeutic benefit for the disease or condition under  
12 treatment. Typically these dosages ranges are as  
13 indicated above for the vaccine formulation. Likewise,  
14 one of ordinary skill in the art can readily determine an  
15 efficacious formulation for delivery of the ratchet  
16 library.

17 In another embodiment of this invention, a ratchet  
18 peptide can be used to identify CTL epitopes within a  
19 protein sequence. Hence, this invention is directed to a  
20 method of constructing a library of related peptides to  
21 provide a ratchet library which comprises identifying a  
22 template peptide; calculating a distribution of amino  
23 acids at each position  $x$  having those amino acids present  
24 in the template peptide at positions  $x$  to  $n-(l-x)$ ,  
25 inclusive, wherein  $l$  is from about 7 to about 25,  $n$  is  
26 from  $l+1$  to about 100, and  $x$  is from 1 to  $l$ ; synthesizing  
27 said ratchet library; and assaying said ratchet library  
28 for the ability to stimulate CTL activity. For example,  
29 the ratchet peptide is constructed and used to immunize  
30 animals, typically though not necessarily mice. The  
31 immunized animals are sacrificed and splenocytes removed  
32 and cultured in vitro with a pool of overlapping  
33 individual peptides that span the ratcheted template. The  
34 activate splenocytes are then tested on target cells  
35 pulsed with 50  $\mu$ M (for example) pooled peptides, and if  
36 any CTL activity is present, the splenocytes are tested on

-21-

1       the individual peptides derived from the region. .If a  
2       single peptide derived from the pool of overlapping  
3       peptides is recognized, a new CTL epitope has been  
4       identified.

5             The examples serve to illustrate the present  
6       invention and are not to be used to limit the scope of the  
7       invention.

-22-

EXAMPLE 1

## Malaria Ratchet Libraries Generate Malaria-Specific CTL

## A. General Methods

Ratchet libraries were synthesized by standard F-moc chemistry using solid phase peptide synthesis with an F-moc RINK MBHA resin [4-(2',4'-dimethoxyphenyl-Fmoc-aminomethyl)-phenoxyacetamido-norleucyl-MBHA resin; MBHA is methylbenzhydrylamine] according to manufacturer's instructions on an ABI Model 433 peptide synthesizer or similar model. The ratchet libraries were synthesized as linear peptides or as branched peptides using a heptalysyl core.

Ratchet libraries were formulated at the indicated concentrations and then used for immunization as microparticles, emulsions or lipopeptides.

Microparticles were prepared according to the water-in-oil-water solvent evaporation method described in U.S. Serial No. 08/263,841, filed June 22, 1994, which is incorporated herein by reference, using polylactide-co-glycolide polymer Resomer RG 505 (Boehringer Ingelheim). Microparticles containing 100 µg of ratchet library were suspended in 0.5 ml phosphate-buffered saline (PBS) for intraperitoneal immunization of mice on days 0, 10 and 20 followed by sacrifice of the animals 7 to 10 days later.

Emulsions were prepared so that the final preparations contained 100 µg of ratchet library and 50 µg Pam<sub>3</sub>Cys-seryl--lysyl-lysyl-lysyl-lysyl (Pam<sub>3</sub>Cys-SKKKK) (SEQ ID NO:12) in a volume of 0.5 mL unless indicated otherwise. To prepare the emulsion, 4 mg of ratchet library was dissolved in 16 mL H<sub>2</sub>O and 240 mg of egg lecithin was dispersed therein by homogenization (Model STD 1 fitted with a 0.25" tubular head, Silverson Machines, East Longmeadow, MA) at 10,000 rpm for 5 min. Pam<sub>3</sub>Cys-SKKK (2 mg) was mixed with 4 g soya oil and then added to the library mixture by homogenization at 10,000

-23-

rpm for 5 min. Further mixing was conducted by ultrasonic pulsation of the emulsion with an ultra sonic probe (Vibra cell, Sonics and Materials, Inc., Danbury, CT). The emulsions (0.5 mL) were injected intraperitoneally into mice on days 0 and 10 followed by sacrifice of the animals 7 to 10 days later.

For lipopeptide ratchet libraries, the peptides of the ratchet library were covalently coupled to Pam<sub>3</sub>Cys (tripalmitoyl-5-glycerol-cysteine) as generally described (Deres et al.) to produce the corresponding lipopeptide ratchet library. The lipopeptide ratchet libraries (100 µg) were suspended in a 0.5 mL of 1% DMSO in glycerol and injected intraperitoneally into mice at day 0. The animals were sacrificed 7-9 days later.

Control mice were injected with 0.5 mL phosphate-buffered saline (PBS) using the corresponding injection schedule as that of the formulated ratchet library.

Upon sacrifice, the spleens were removed and splenocytes were pooled and cultured in vitro with the indicated peptide at a concentration of 1 µg/mL for one week to produce activated splenocytes.

CTL assays were then conducted according to the method of McDonald et al. (1980) as briefly described below.

#### B. Malarial Ratchet Libraries

A ratchet library was designed from a 20 amino acid sequence of residues 247-266 from the circumsporozoite protein of *Plasmodium berghei* (CS 247-266). This sequence contains the nonamer CTL epitope designated as CS 252-260 [Eberl et al. (1993) Int. Immunol. 5:1489-1492].

Malarial ratchet library 1 from the template peptide of SEQ ID NO:1 (Fig.1B, bottom panel) was prepared in linear form and formulated in microparticles at a final concentration of 200 µg/mL. Three BALB/C mice were immunized with 0.5 mL per injection as described above.

-24-

1 Pooled splenocytes were cultured with *Plasmodium berghei*  
2 CS peptide 252-260.

3 CS 252-260-specific CTL activity was then assayed.  
4 Briefly, H-2<sup>d</sup> target cells (mouse mastocytoma cell line  
5 P815 or A20.1) were <sup>51</sup>Cr labeled for one h, washed and then  
6 incubated for one h in the presence of CS 252-260 peptide  
7 at 50  $\mu$ M, in the presence of an unrelated control peptide,  
8 influenza nucleoprotein peptide NP 147-155 at 50  $\mu$ M or in  
9 media (i.e., in the absence of a peptide antigen). CTL  
10 activity was determined by incubating these target cells  
11 with activated splenocytes (effector cells) for 4 hours in  
12 round-bottomed 96-well plates at a range of  
13 effector:target (E:T) ratios of 100:1, 50:1, 25:1 and  
14 12.5:1 and measuring the release of <sup>51</sup>Cr. Cell lysis was  
15 calculated as per cent target cell lysis from the formula  
16  $(E-M/T-M) \times 100$ , where E = experimental <sup>51</sup>Cr release (cpm);  
17 M = <sup>51</sup>Cr release in presence of culture medium; and T =  
18 total <sup>51</sup>Cr released by 10% Triton X-100®.

19 The results are shown in Fig. 2 and indicate that  
20 significant CTL lysis was elicited in an MHC-restricted  
21 manner, since the K<sup>d</sup>-restricted CS 252-260 peptide epitope  
22 sensitized target cells incubated with that peptide and  
23 not with any of the controls. Significant CTL lysis  
24 occurs if there is greater than 10% lysis above the  
25 control level of lysis at the highest E:T ratio.  
26 Splenocytes from control mice did not elicit specific CTL  
27 activity in any experiment.

28 Malaria ratchet library 2 (Fig. 1B, bottom panel) was  
29 synthesized as branched octameric peptides on a heptalysyl  
30 core and formulated in microparticles at a final  
31 concentration of 2 mg/mL and injected into mice as  
32 described above. Splenocytes were incubated and CTL  
33 activity was assayed as described above using E:T ratios  
34 of 100:1, 50:1 25:1 and 12.5:1. The results are shown in  
35 Fig. 3.

-25-

1 Malaria ratchet library 2 was formulated and injected  
2 into mice as described in the preceding paragraph to  
3 determine CTL activity against the 20 residue CS 247-266  
4 peptide. Splenocytes were incubated with CS 252-260  
5 peptide and CTL activity was determined as above, except  
6 that the target cells were incubated in the presence of 50  
7  $\mu$ M of CS 252-260 peptide, 50  $\mu$ M of CS 247-266 peptide or  
8 in media using E:T ratios of 100:1, 50:1 25:1 and 12.5:1.  
9 The results are shown in Fig. 4.

10 Malaria ratchet library 1 was formulated and injected  
11 as a lipopeptide except that the lipopeptide ratchet  
12 library was formulated at a concentration of 20  $\mu$ g/mL.  
13 Splenocytes were incubated CS 252-160 and CTL activity was  
14 assayed as described above using E:T ratios of 100:1, 50:1  
15 25:1 and 12.5:1. The results are shown in Fig. 5.

16 Malaria ratchet library 1 was formulated and injected  
17 as a lipopeptide ratchet library at a concentration of 200  
18  $\mu$ g/mL. Splenocytes were incubated with CS 252-260 and CTL  
19 activity was as described above to determine CTL activity  
20 against a self peptide. For the CTL activity  
21 determination, the target cells were incubated in the  
22 presence of 50  $\mu$ M of CS 252-260 peptide, 50  $\mu$ M of a K<sup>d</sup>-  
23 restricted self peptide (SYFPEITHI; SEQ ID NO:13) or in  
24 media using E:T ratios of 100:1, 50:1 25:1 and 12.5:1.  
25 The results are shown in Fig. 6.

26 These malaria ratchet libraries elicit malaria-  
27 specific CTL at immunogen doses ranging from 10  $\mu$ g to 1 mg  
28 (Figs. 2-5). The CTL activity is MHC restricted and is  
29 elicited when presented with the processed epitope as a  
30 longer peptide (Fig. 4). CTL activity was absent when  
31 target cells were pulsed with a K<sup>d</sup> restricted self-peptide  
32 to test for any auto-immune reaction. No significant  
33 lysis was seen on the targets pulsed with the self  
34 peptide.

-26-

EXAMPLE 2

## MHC-Binding Motif Restricted Ratchet Library

Fig. 7, bottom panel, illustrates an MHC-restricted malaria ratchet library from the template peptide of SEQ ID NO:1. This library is constructed from malaria ratchet library 1 of Fig. 1B (also top panel of Fig. 7) by replacing those positions which contain anchor residues for the K<sup>b</sup>, D<sup>b</sup>, K<sup>d</sup>, and L<sup>d</sup> molecules (see below) with an equal proportion of the anchor residues at the position in question. The anchor residues are those amino acids which have been identified as necessary for binding to the MHC class I molecule for the given haplotype. The anchor residues for the indicated haplotypes are shown in the middle panel of Fig. 7 and the MHC-restricted malaria ratchet library is shown in the bottom panel of Fig. 7.

At each of these positions (i.e., positions 2, 5, 8 and 9), the ratchet incorporates only those anchor amino acids shown in the middle panel. Thus, position 2 contains 50% tyrosine Y and 50 % proline; position 5 contains 33% asparagine, 33% tyrosine and 33% phenylalanine; position 8 contains 100% leucine and position 9 contains 25% isoleucine, 25% leucine, 25% phenylalanine and 25% methionine. Thus, this ratchet has the anchor residues involved in binding MHC class I molecules and stimulating CTL.

-27-

EXAMPLE 3**Ratchet Libraries Can Accomodate  
Antigenic Diversity**

An HIV ratchet library was constructed from the SSAL (Fig. 8A, bottom panel) of template peptide SEQ ID NO:2 and the additional 14 HIV-1 sequences (Fig. 8A) from a 35 amino acid sequence of the HIV-1 gp120 V3 loop region, the principle neutralizing domain known to have extensive sequence variability. This region contains a D<sup>4</sup> restricted CTL epitope at amino acid positions 318-326. The antigenic diversity of this region is accommodated by taking 15 HIV-1 consensus sequences including the sequence HIV-MVP5188 and constructing an SSAL library where the identity and ratio of amino acids at each position is determined by the relative prevalence of amino acids in those 16 sequences. Next, it is the SSAL library which is "ratcheted" to yield the HIV-1 ratchet library shown in the bottom panel of Fig. 8.

-28-

EXAMPLE 4Peptide Ratchet Libraries from  
Different Length Template Peptides

To examine the effect of template peptide length on the efficacy of CTL induction via ratchet libraries, three HIV-1 gag peptide linear ratchet libraries (Figs. 9-11, bottom panel) containing a mouse HIV CTL epitope were synthesized using template peptides of lengths 100, 40 or 20 amino acids of the gag sequence as shown in Figs. 9-11, respectively, and designated by SEQ ID NOS:3-5, respectively.

These libraries were formulated with 100  $\mu$ g in 0.5 mL as microparticles, emulsions or lipopeptides and injected as described in Example 1 with the following modifications: The immunized mice were C57BL/6 mice. Activated splenocytes were prepared by culturing with 1  $\mu$ g/mL HIV gag peptide 390-398 [Elvin et al. (1993) J. Immunol. Methods 158:161-171]. The target cells were EL4 and were incubated with 50  $\mu$ M HIV gag peptide 390-398. The results are presented in Table 1. Fig. 12 show the specific cell lysis results for the emulsion formulation of the ratchet library from the 40 amino acid template.

-29-

Table 1

## CTL Response from HIV-1 Ratchet Libraries

Formulation <sup>a</sup>	Pam <sub>3</sub> Cys <sup>b</sup>	Template Peptide Length <sup>c</sup>		
		20	40	100
Lipopeptide	+ <sup>d</sup>	+	±	++
Emulsion	+	±	+++	-
Microparticle	-	-	-	-

<sup>a</sup> The formulations are described in Example 1 and the experimental protocol in Example 4.

<sup>b</sup> Pam<sub>3</sub>Cys is covalently bound to the ratchet library for the lipopeptide preparation but not in the emulsion preparation.

<sup>c</sup> The template peptide of the indicated amino acid length was ratcheted to 9-mers as described in Example 4.

<sup>d</sup> The symbols are as follows: +++, ++ and + mean significant specific target cell lysis in the indicated relative amounts with +++ as the most lysis; ±, inconclusive amount of specific cell lysis; -, no significant specific cell lysis.

-30-

EXAMPLE 5

## Mucin Ratchet Libraries

Fig 13. (bottom panel) provides a mucin ratchet library constructed from the template peptide of SEQ ID NO:7 or SEQ ID NO:8. Mucin is a large, heavily glycosylated molecule expressed and secreted by ductal epithelial cells and tumors. Mucin consists of multiple copies of a 20 amino acid tandem repeat (SEQ ID NO:6) which appears to elicit non-MHC restricted CTL responses.

Because the 20-mer repeat will not contain every possible nonomer when used as a template, due to end effects in the ratchet, an alternative approach was used to generate all possible nonomers of the repeating 20-mer peptide. In this case the last eight carboxy-terminal amino acids of the 20-mer repeat peptide were placed at its amino terminus to yield a template peptide of 28 amino acids before calculation of the ratchet (Fig. 13, middle panel). Alternatively, the first eight amino-terminal amino acids were placed at its carboxyl terminus to yield a template peptide of 28 amino acids before calculation of the ratchet (Fig. 13, middle panel). With either method the calculated mucin ratchet library is the same.

-31-

EXAMPLE 6

## Mutant p53 Ratchet Libraries

The protein p53 is a tumor supressor which fails to function effectively when mutated. More than 50% of human tumors contain cells which express a mutant form of p53, due to one or more point mutations in the protein. Class I mutations in p53 affect residues that directly contact DNA and include the residues lysine at postion 120, serine at position 241, arginine at position 248, arginine at position 273, alanine at position 276, cysteine at position 277 and arginine at positon 283. In this group, the mutations of arginine at positions 248 and 273 appear most frequently. Class II mutants affect residues that do not contact DNA but rather appear to have a role in stabilizing protein structure and include mutations of arginine at positions 175 and 249.

This example provides four peptide ratchet libraries containing four hot spots of mutation of p53.

The four ratchets are designed around mutation hot spots in the protein: (1) template peptide of amino acids 124-151 (SEQ ID NO:9) as a 10-mer ratchet library (Fig. 14, bottom panel); (2) template peptide of amino acids 166-187 as a 9-mer ratchet library; (3) template peptide of amino acids 228-256 as a 9-mer ratchet library; and (4) template peptide of amino acids 264-289 as a 9-mer ratchet library.

The ratchet library from the 166-187 template is especially useful for treating colon cancer since this mutation is frquently encountered with this malignancy.

-32-

EXAMPLE 7

## Influenza Ratchet Libraries

Existing influenza vaccines are complex to design and manufacture as the prevalent strain of influenza can rapidly change and a vaccine designed to stimulate antibodies against influenza A of strain one may not be effective in eliciting cross-reactive immunity against strain two. The CTL response has been shown to be effective against influenza in animal models and in humans, and the addition of an influenza specific CTL component to existing vaccines, or a CTL inducing vaccine alone, would dramatically broaden protection against many strains of influenza.

To provide a vaccine capable of stimulating a CTL response against influenza, known CTL epitopes can be ratcheted. Fig. 15 (bottom panel) shows influenza ratchet library 1 from a 25-mer template peptide of residues 139-163 (SEQ ID NO:10) of influenza A A/34/PR8 nucleoprotein. This library encompasses the known K<sup>d</sup>-restricted epitope of residues 147-155. Fig. 16 (bottom panel) shows influenza ratchet library 2 constructed from the template peptide of SEQ ID NO:11 which is a linkage of 3 CTL epitopes in order from N to C terminus of residues 50-58, 147-155 and 366-374 of the nucleoprotein.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Kuebler, Peter J.  
Nixon, Douglas F.
- (ii) TITLE OF INVENTION: Peptide Ratchets for Vaccines  
and Therapeutics
- (iii) NUMBER OF SEQUENCES: 13
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: M. Lisa Wilson
  - (B) STREET: 25 Davids Drive
  - (C) CITY: Hauppauge
  - (D) STATE: NY
  - (E) COUNTRY: USA
  - (F) ZIP: 11788
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version  
#1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Wilson, M. Lisa
  - (B) REGISTRATION NUMBER: 34,045
  - (C) REFERENCE/DOCKET NUMBER: 2012
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (516)273-2828
  - (B) TELEFAX: (516)273-1717

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

-34-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

Asn Asn Asn Asp Asp Ser Tyr Ile Pro Ser Ala Glu
 1           5           10
Lys Ile Leu Glu Phe Val Lys Gln
      15           20

```

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile
 1           5           10
His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly
      15           20
Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys
 25           30           35

```

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 100 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr
 1           5           10
Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu Gln
      15           20
Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr
 25           30           35
Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr
      40           45
Ile Leu Lys Ala Leu Gly Pro Ala Ala Thr Leu Glu
 50           55           60
Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro
      65           70
Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
      75           80
Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg
 85           90           95

```

-35-

Gly Asn Phe Leu  
100

## (2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 40 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Glu	Met	Met	Thr	Ala	Cys	Gln	Gly	Val	Gly	Gly	Pro
1				5					10		
Gly	His	Lys	Ala	Arg	Val	Leu	Ala	Glu	Ala	Met	Ser
	15				20						
Gln	Val	Thr	Asn	Ser	Ala	Thr	Ile	Met	Met	Gln	Arg
25			30							35	
Gly	Asn	Phe	Leu								
			40								

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 20 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Glu	Ala	Met	Ser	Gln	Val	Thr	Asn	Ser	Ala	Thr	Ile
1				5					10		
Met	Met	Gln	Arg	Gly	Asn	Phe	Leu				
	15				20						

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 20 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Pro	Asp	Thr	Arg	Pro	Ala	Pro	Gly	Ser	Thr	Ala	Pro
1				5					10		

-36-

Pro Ala His Gly Val Thr Ser Ala  
           15                          20

## (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 28 amino acids  
     (B) TYPE: amino acid  
     (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg  
   1                          5                          10  
 Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala His Gly  
           15                          20  
 Val Thr Ser Ala  
   25

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 28 amino acids  
     (B) TYPE: amino acid  
     (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Pro Asp Thr Arg Pro Ala Pro Gly Ser Thr Ala Pro  
   1                          5                          10  
 Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg  
           15                          20  
 Pro Ala Pro Gly  
   25

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 28 amino acids  
     (B) TYPE: amino acid  
     (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Cys Thr Tyr Ser Pro Ala Leu Asn Lys Met Phe Cys  
   1                          5                          10

-37-

Gln Leu Ala Lys Thr Cys Pro Val Gln Leu Trp Val  
                   15                                  20  
 Asp Ser Thr Pro  
                   25

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Trp His Ser Asn Leu Asn Asp Ala Thr Tyr Gln Arg  
   1                                  5                                  10  
 Thr Arg Ala Leu Val Arg Thr Gly Met Asp Pro Arg  
                   15                                  20  
 Met  
                   25

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Ser Asp Tyr Glu Gly Arg Leu Ile Thr Tyr Gln Arg  
   1                                  5                                  10  
 Thr Arg Ala Leu Val Ala Ser Asn Glu Asn Met Glu  
                   15                                  20  
 Thr Met  
                   25

## (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 6 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

-38-

## (ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 1

(D) OTHER INFORMATION: /note= "N-terminal  
tripalmitoyl-5-glycerol"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:.

Cys Ser Lys Lys Lys Lys  
1 5

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ser Tyr Phe Pro Glu Ile Thr His Ile  
1 5

-39-

## We Claim:

1. A ratchet library of peptides comprising at least one immunostimulatory cytotoxic T lymphocyte (CTL) epitope wherein said peptides are of length  $l$ ; the sequences of said peptides in said library being determined from a template peptide of length from  $l+1$  to  $n$  amino acids such that each position  $x$  in the library has all the amino acids present in said template peptide at positions  $x$  to  $n-(l-x)$ , inclusive; the ratio of amino acids at each position  $x$  being determined by the relative prevalence of amino acids at that position  $x$ ; and wherein  $l$  is from about 7 to about 25 amino acids,  $n$  is from  $l+1$  to about 100, and  $x$  is from 1 to  $l$ .

2. The library of Claim 1 wherein  $n$  is from  $l+1$  to about 75.

3. The library of Claim 1 wherein  $n$  is from  $l+1$  to about 50.

4. The library of Claim 1 wherein  $l$  is from 8 to 10.

5. The library of Claim 1 wherein  $l$  is 9.

6. The library of Claim 1 wherein if a position  $x$  is identified as part of an MHC-binding motif of a CTL epitope, then that position  $x$  is fixed as one or more amino acids of said MHC-binding motif in an equimolar ratio.

7. The library of any one of Claims 1 to 6 wherein said CTL epitope is from a virus, bacterium, parasite, tumor antigen, allergen or other protein antigen.

8. The library of Claim 1 to 6 wherein said CTL epitope is from a melanoma protein including MAGE-1, -2, and -3; a renal cell carcinoma protein; a colon carcinoma protein; a prostate cancer protein (malignant or benign) including PSA; tyrosinase; an oncogene such as HER-2/neu proto-oncogene; ras; MUC1; p53; p16; TL; an HIV-1 or HIV-2 protein including envelope, gag, pol, nef, tat, rev, vpx or vpu; an HTLV I or II protein including envelope, gag, pol, pX or TAX; lymphocytic choriomeningitis virus of

-40-

mice; influenza A, B or C including PB1, PB2, PA, NS1, M1, NP or HA; an Epstein-Barr virus protein including TETA, EENL, EBNA3, EBNA1 or LMP; respiratory syncytia virus; hepatitis B virus; hepatitis C virus; herpes simplex virus; cytomegalovirus; a parainfluenza virus 1 protein including hemagglutinin, neuraminidase, phosphoprotein or nucleoprotein; intracisternal A particle gag; bovine leukemia virus; papilloma viruses; a malaria protein including proteins from *P. falciparum*, *P. berghei*, *P. ovale*, *P. vivax*, *P. malaria*; *Histoplasma capsulatum*; *Listeria*; *Toxoplasmosis*; *Trypanosoma cruzi*; *Yersinia*; *M. tuberculi*; *M. lepri*; *Pneumocystis carinii*; Kaposi's sarcoma or frameshift sequences.

9. The library of Claim 1 wherein said template peptide is any one of SEQ ID NOS: 1-11.

10. The library of any one of Claims 1-6 or 9 wherein said peptides of said library have a covalently attached N-terminal tripalmitoyl-5-glycerol-cysteine moiety.

11. The library of any one of Claims 1-6 or 9 wherein said peptides are linked to a branched core sequence, are polymerized or are conjugated to a carrier molecule.

12. A pharmaceutical or vaccine composition comprising the library of any one of Claims 1-6 or 9, or a pharmaceutically acceptable salt thereof, and a pharmaceutically acceptable carrier.

13. The composition of Claim 12 wherein said composition is an emulsion or a microparticle formulation.

14. The composition of Claim 12 wherein said formulation also comprises tripalmitoyl-5-glycerolcysteine or a derivative thereof.

15. A pharmaceutical or vaccine composition comprising the library of Claim 7, or a pharmaceutically acceptable salt thereof, and a pharmaceutically acceptable carrier.

-41-

16. The composition of Claim 15 wherein said composition is an emulsion or a microparticle formulation.

17. The composition of Claim 15 wherein said formulation also comprises tripalmitoyl-5-glycerolcyteine or a derivative thereof.

18. A method of treating or preventing a disease or a malignancy which comprises administering an amount of said composition of Claim 12 to a mammal effective to stimulate a CTL response against said disease or said malignancy associated with said CTL epitope present in said library.

19. A method of treating or preventing a disease or a malignancy which comprises administering an amount of said composition of Claim 15 to a mammal effective to stimulate a CTL response against said disease or said malignancy associated with said CTL epitope present in said library.

20. A method of constructing a library of related peptides to provide a ratchet library which comprises identifying a template peptide; calculating a distribution of amino acids at each position  $x$  having those amino acids present in the template peptide at positions  $x$  to  $n-(1-x)$ , inclusive, wherein  $l$  is from about 7 to about 25,  $n$  is from  $l+1$  to about 100, and  $x$  is from 1 to  $l$ ; and synthesizing said ratchet library.

[illegible]

Figure 1A

2/18

## Sequence Alignment by Position

	1	2	3	4	5	6	7	8	9
N	N	N	N	D	D	S	Y	I	P
N	N	D	D	D	S	Y	I	P	S
N	D	D	S	Y	I	P	S	A	E
D	S	Y	I	P	S	A	E	K	I
D	S	I	P	S	A	E	K	I	L
S	Y	P	S	A	E	K	I	L	E
Y	I	P	S	A	E	K	I	L	F
P	S	A	E	K	I	L	E	F	V
S	A	E	K	I	L	E	F	V	K
A	E	K	I	L	E	F	V	K	Q

## AA Distribution by Position

	1	2	3	4	5	6	7	8	9
AA residue									
N	3	2	1		1				
D	2	2	2	2	2	2	1	1	1
S	2	2	2	2	2	2	1	1	1
Y	1	1	1	1	1	1	1	2	1
I	1	1	2	2	2	2	2	1	1
P	1	1	1	1	1	1	1	1	1
A	1	1	1	1	1	1	1	1	1
E	1	1	1	1	2	2	2	2	2
K		1	1	1	1	1	1	2	2
L				1	1	1	1	1	1
F						1	1	1	1
V							1	1	1
Q									1

## AA Percentage by Position

	1	2	3	4	5	6	7	8	9
AA residue									
N	25	17	8						
D	17	17	17	17	8				8
S	17	17	17	17	17	17	8	8	
Y	8	8	8	8	8	8	8		
I	8	8	17	17	17	17	17	17	8
P	8	8	8	8	8	8	8	8	8
A	8	8	8	8	8	8	8	8	8
E	8	8	8	8	17	17	17	17	17
K		8	8	8	8	8	8	17	17
L				8	8	8	8	8	8
F						8	8	8	8
V							8	8	8
Q									8

Figure 1B

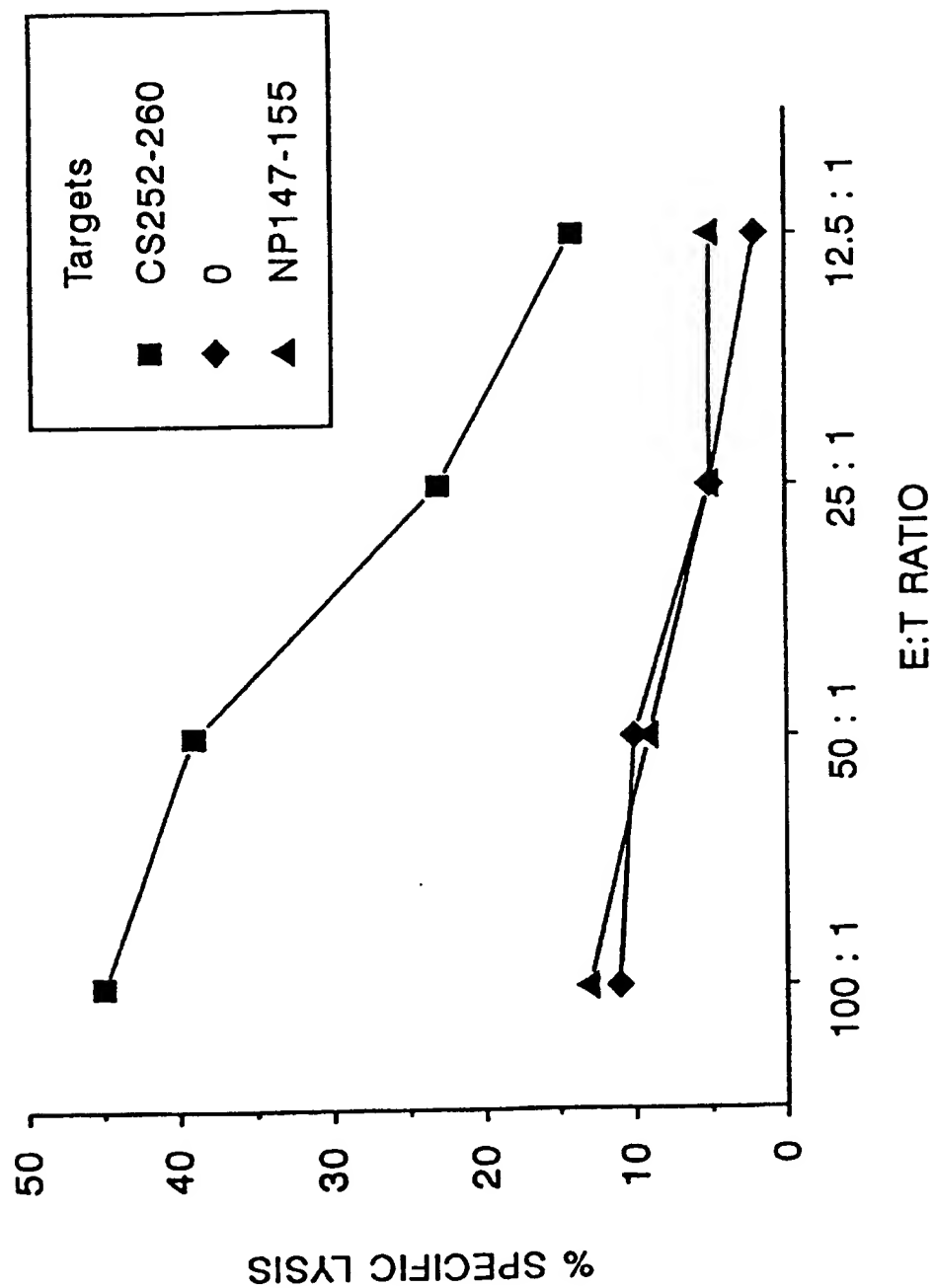


Figure 2

4/18

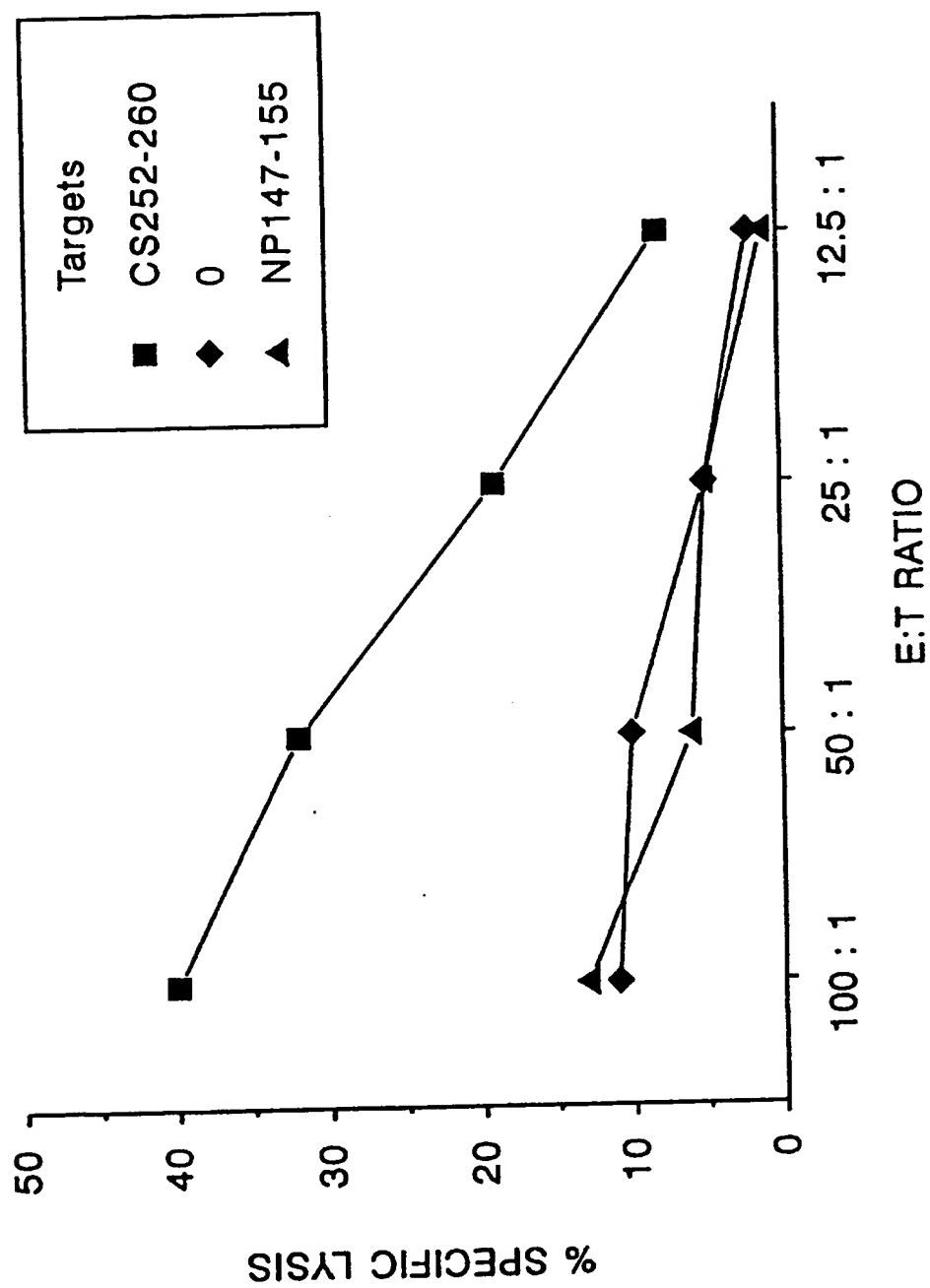


Figure 3

5/18

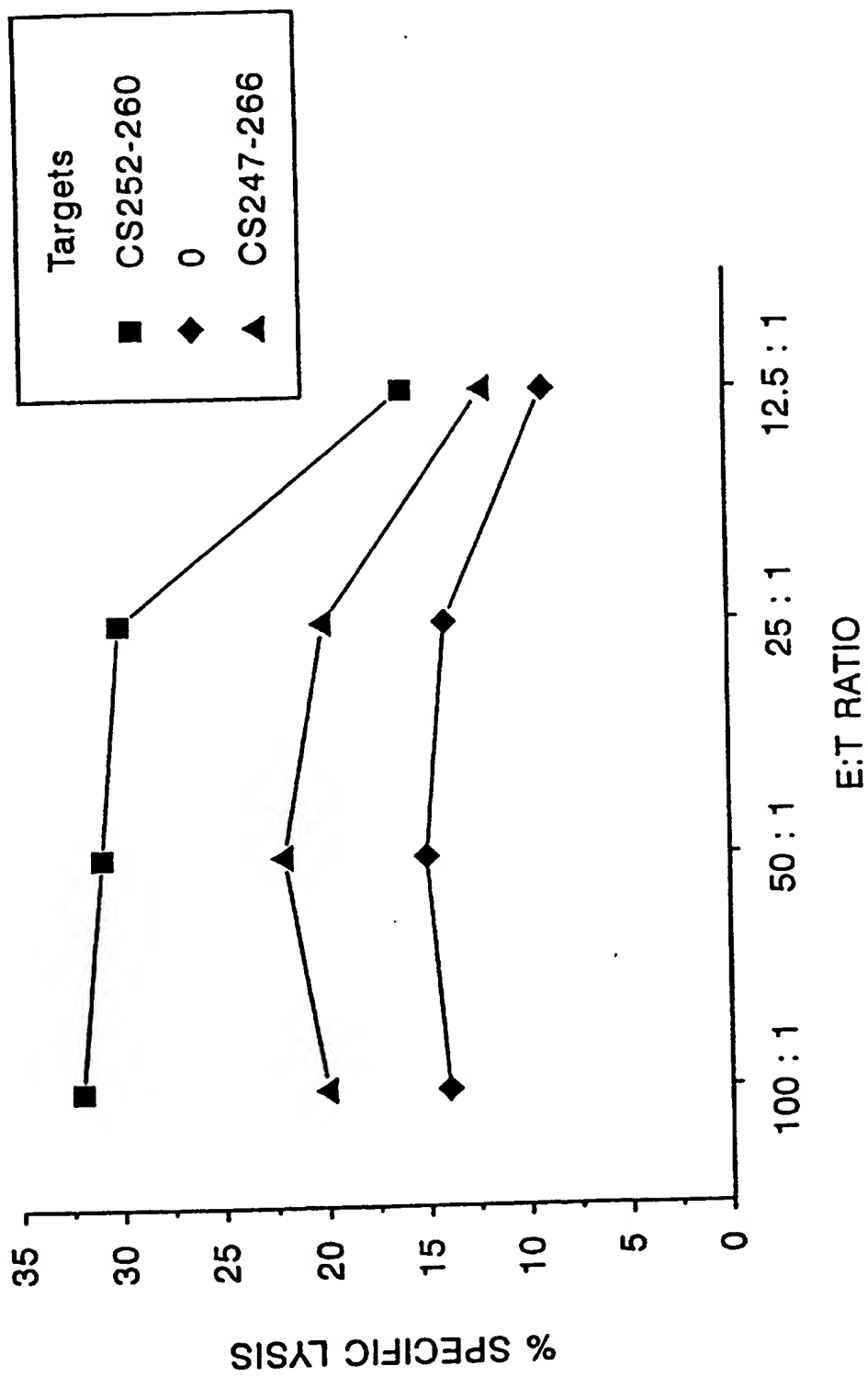


Figure 4

6/18

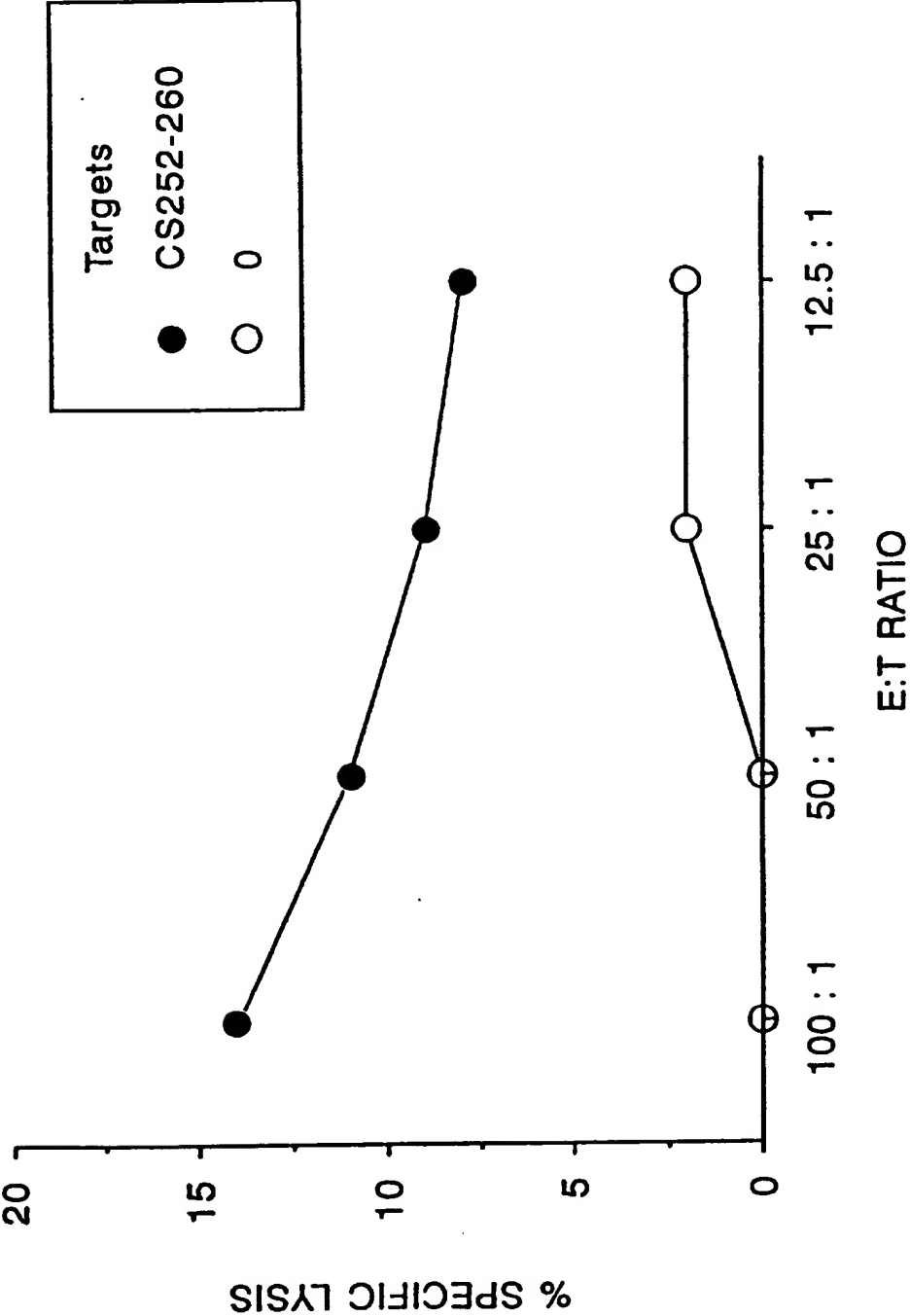


Figure 5

7/18

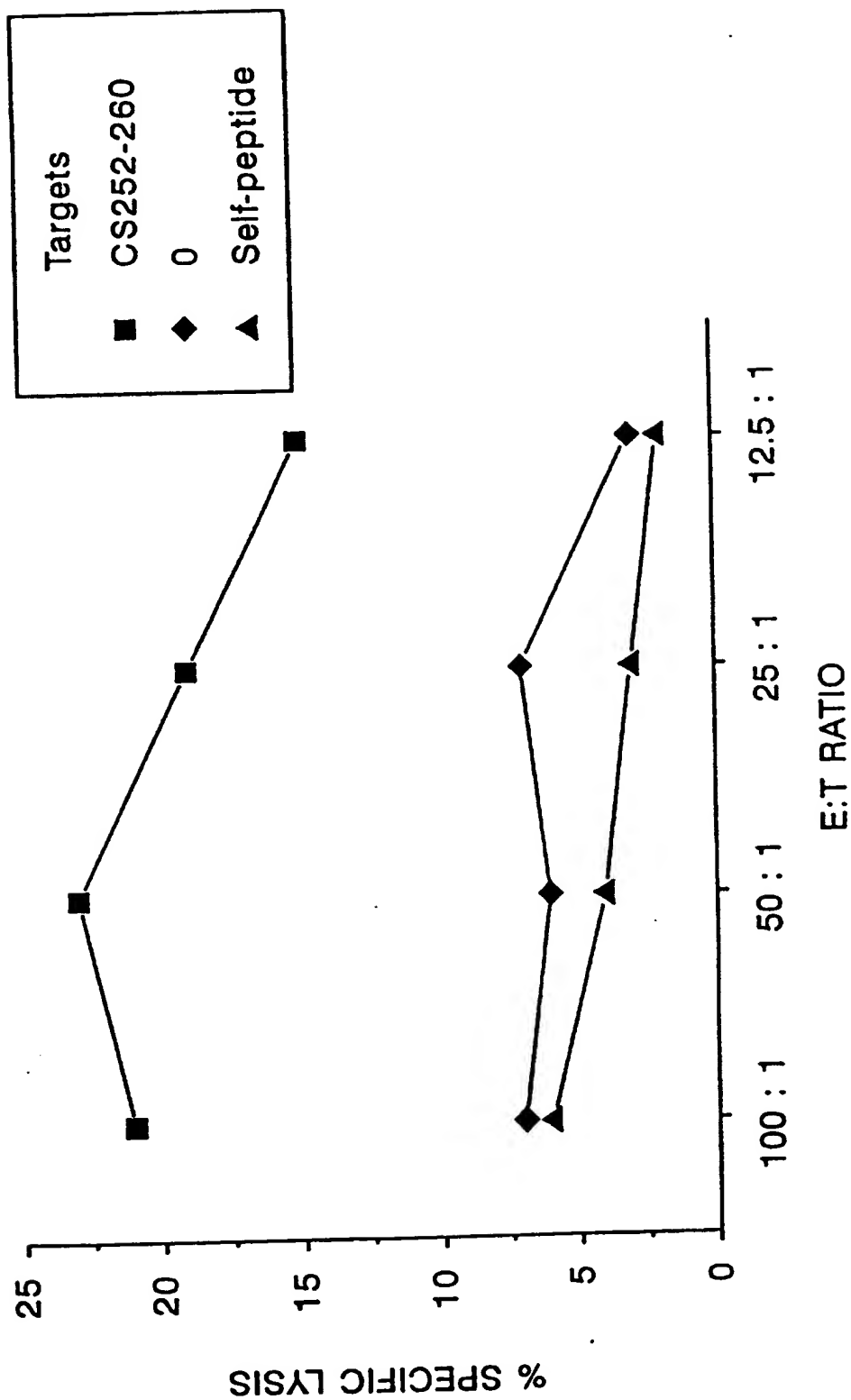


Figure 6

8/18

AA residue	AA Percentage by Position								
	1	2	3	4	5	6	7	8	9
N	25	17	8						
D	17	17	17	17	8				
S	17	17	17	17	17	17	8	8	8
Y	8	8	8	8	8	8	8		
I	8	8	17	17	17	17	17	17	8
P	8	8	8	8	8	8	8	8	8
A	8	8	8	8	8	8	8	8	8
E	8	8	8	8	17	17	17	17	17
K		8	8	8	8	8	8	17	17
L				8	8	8	8	8	8
F						8	8	8	8
V							8	8	8
Q									8

AA Restrictions by Position	
Haplotype	
K <sup>d</sup>	Y I
	L
D <sup>b</sup>	N M
K <sup>b</sup>	F L
	Y
L <sup>d</sup>	P M
	L
	F

AA residue	Restricted AA Percentage by Position								
	1	2	3	4	5	6	7	8	9
N	25		8		33				
D	17		17	17					
S	17		17	17		17	8		
Y	8	50	8	8	33	8	8		
I	8		17	17		17	17		25
P	8	50	8	8		8	8		
A	8		8	8		8	8		
E	8		8	8		17	17		
K			8	8		8	8		
L				8		8	8	100	25
F					33	8	8		25
V							8		
Q									
M									25

Figure 7

CONSENSUS	B	C	T	R	P	N	N	T	R	K	S	I	H	I	G	P	G	R	A	F	Y	T	T	G	E	I	G	D	I	R	Q	A	H	C
1	c	t	r	p	n	n	n	t	r	k	s	i	h	i	g	p	g	r	a	f	y	t	t	g	d	i	g	d	i	r	q	a	h	c
2	c	t	r	p	n	n	n	t	r	k	s	i	h	i	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
3	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
4	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
5	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
6	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
7	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
8	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
9	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
10	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
11	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
12	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
13	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
14	c	t	r	p	n	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c
HIVMVP5180		c	t	r	p	n	n	t	r	k	s	i	h	s	g	p	g	r	a	f	y	a	t	g	d	i	g	d	i	r	q	a	h	c

C T R P N N T R K S I H I G P G R A F Y T T G E I I G D I R Q A H C  
C15 T13R15P:4 N5N10N13T11 R12 K6 S7 I5 H8 I13 G15P12G14 R9 A10 F11 Y14 T7 T15 G8 E2 I10 I8 G9 D8 I9 R9 Q7 A9 H7 C9  
I2 E1 S2 K2 K1 M1 S1 T2 R3 T3 R2P1 Q1 M1 Q6 V2 L3 L1 A2 K1 D4 R3 T1 D3 I5 R3 C4 K2 H3 Y2 R1  
F1 D1 A1 I1 K1 Q6 A1 P1 S2 T1 L2 M1 Q1 Y1 K2 N1 K2 T1 G4 Y1 N1 G2 K1 A5 R1 S1  
Y6 Q1 K1 V1 I1 G2 A1 T2 S1 S1 T1 R1 E1 Q1 K1 S2 I1 T1 S1 P1 C5  
G1 I1 E1 Q1 P2 N1 L1 S2 N1 R3 I4

**Figure 8A**

10/16

AA Distribution by Position									
AMINO ACID	1	2	3	4	5	6	7	8	9
C	15							5	9
T	59	59	46	47	47	47	47	47	36
R	48	48	48	36	45	45	46	46	45
P	29	29	29	29	16	16	16	16	16
I	48	49	42	51	51	51	50	50	49
K	17	17	17	17	18	18	16	15	14
E	5	5	5	5	4	4	4	4	3
N	30	31	30	30	30	25	15	2	2
F	12	12	12	12	12	11	11	11	11
Y	22	23	23	23	23	17	17	17	17
G	44	53	53	55	55	54	54	54	54
H	6	6	6	6	6	6	11	11	11
L	7	7	7	7	7	7	7	7	7
Q	17	17	17	17	21	21	20	20	20
S	17	17	17	17	18	16	16	16	16
D	6	9	18	18	18	18	17	17	17
A	17	17	17	17	17	17	26	25	25
V	3	3	3	3	3	3	3	3	3
M	3	3	3	3	3	3	3	3	2

AA Percentage by Position									
AMINO ACID	1	2	3	4	5	6	7	8	9
C	4							1	3
T	15	15	12	12	12	13	13	13	10
R	12	12	12	9	12	12	12	13	13
P	7	7	7	7	4	4	4	4	5
I	12	12	11	13	13	13	13	13	14
K	4	4	4	4	4	4	4	4	4
E	1	1	1	1	1	1	1	1	1
N	7	7	8	8	8	7	4	1	1
F	3	3	3	3	3	3	3	3	3
Y	5	5	6	6	6	5	5	5	5
G	11	13	13	14	14	14	14	15	15
H	2	2	2	2	2	2	3	3	3
L	2	2	2	2	2	2	2	2	2
Q	4	4	4	4	5	5	5	5	6
S	4	4	4	4	4	4	4	4	5
D	2	2	5	5	5	5	5	5	5
A	5	5	5	5	5	5	7	7	7
V	1	1	1	1	1	1	1	1	1
M	1	1	1	1	1	1	1	1	1

Figure 8B

11/18

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25  
 I R Q G P K E P F R D Y V D R F Y K T L R A E Q A  
 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50  
 S Q E V K N W M T E T L L V Q N A N P D C K T I L  
 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75  
 K A L G P A A T L E E M M T A C Q G V G G P G H K  
 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100  
 A R V L A E A M S Q V T N S A T I M M Q R G N F L

AA Distribution by Position		1	2	3	4	5	6	7	8	9
AMINO ACID										
I		3	2	2	2	2	2	2	2	2
R		5	5	4	4	5	5	5	5	5
Q		6	6	6	6	6	6	6	6	6
G		6	6	6	6	5	6	6	6	6
P		5	5	5	5	5	4	4	4	3
K		6	6	6	6	6	6	5	5	5
E		7	7	7	7	7	7	7	6	6
F		2	2	2	2	2	2	2	3	3
D		3	3	3	3	3	3	3	3	3
Y		2	2	2	2	2	2	2	2	2
V		6	6	6	6	6	6	6	6	6
T		8	8	8	8	8	8	8	8	8
L		7	7	7	7	7	7	7	7	8
A		11	11	11	11	11	11	11	11	11
N		4	4	4	4	4	4	5	5	5
S		3	3	3	3	3	3	3	3	3
W		1	1	1	1	1	1	1	1	1
M		4	5	6	6	6	6	6	6	6
C		2	2	2	2	2	2	2	2	2
H		1	1	1	1	1	1	1	1	1

AA Percentage by Position		1	2	3	4	5	6	7	8	9
AMINO ACID										
I		3	2	2	2	2	2	2	2	2
R		5	5	4	4	5	5	5	5	5
Q		7	7	7	7	7	7	7	7	7
G		7	7	7	7	7	7	7	7	7
P		5	5	5	5	5	4	4	4	3
K		7	7	7	7	7	7	5	5	5
E		8	8	8	8	8	8	8	7	7
F		2	2	2	2	2	2	2	3	3
D		3	3	3	3	3	3	3	3	3
Y		2	2	2	2	2	2	2	2	2
V		7	7	7	7	7	7	7	7	7
T		9	9	9	9	9	9	9	9	9
L		8	8	8	8	8	8	8	8	9
A		12	12	12	12	12	12	12	12	12
N		4	4	4	4	4	4	5	5	5
S		3	3	3	3	3	3	3	3	3
W		1	1	1	1	1	1	1	1	1
M		4	5	7	7	7	7	7	7	7
C		2	2	2	2	2	2	2	2	2
H		1	1	1	1	1	1	1	1	1

Figure 9

12/18

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20  
 E M M T A C Q G V G G P G H K A R V L A

21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40  
 E A M S Q V T N S A T I M M Q R G N F L

AA Distribution by Position									
AMINO ACID	1	2	3	4	5	6	7	8	9
I	1	1	1	1	1	1	1	1	1
R	1	1	1	1	2	2	2	2	2
Q	2	2	2	3	3	3	3	2	2
G	4	4	4	4	4	5	5	5	4
P	1	1	1	1	1	1	1	1	1
K	1	1	1	1	1	1	1	1	1
E	2	1	1	1	1	1	1	1	1
V	3	3	3	3	3	3	3	3	3
T	3	3	3	3	2	2	2	2	2
L	1	1	1	1	1	1	1	1	2
A	5	5	5	5	5	4	4	4	4
N	1	1	1	1	1	1	2	2	2
S	2	2	2	2	2	2	2	2	2
M	3	4	4	3	3	3	3	3	3
C	1	1	1	1	1	1			
H	1	1	1	1	1	1	1	1	1
F								1	1

AA Percentage by Position									
AMINO ACID	1	2	3	4	5	6	7	8	9
I	3	3	3	3	3	3	3	3	3
R	3	3	3	3	6	6	6	6	6
Q	6	6	6	9	9	9	9	6	6
G	13	13	13	13	13	16	16	16	13
P	3	3	3	3	3	3	3	3	3
K	3	3	3	3	3	3	3	3	3
E	6	3	3	3	3	3	3	3	3
V	9	9	9	9	9	9	9	9	9
T	9	9	9	9	6	6	6	6	6
L	3	3	3	3	3	3	3	3	6
A	16	16	16	16	16	13	13	13	13
N	3	3	3	3	3	3	6	6	6
S	6	6	6	6	6	6	6	6	6
M	9	13	13	9	9	9	9	9	9
C	3	3	3	3	3	3			
H	3	3	3	3	3	3	3	3	3
F								3	3

Figure 10

13/18

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
E	A	M	S	Q	V	T	N	S	A	T	I	M	M	Q	R	G	N	F	L

AMINO ACID	AMINO ACID POSITION#								
	1	2	3	4	5	6	7	8	9
I	1	1	1	1	1	1	1	1	1
R					1	1	1	1	1
Q	1	1	1	2	2	1	1	1	1
G						1	1	1	1
E	1								
V	1	1	1	1	1	1			
T	2	2	2	2	2	2	2	1	1
L									1
A	2	2	2	2	2	2	2	2	2
N	1	1	1	1	1	1	2	2	1
S	2	2	2	2	1	1	1	1	1
M	1	2	2	1	1	1	1	1	1
F								1	1

AMINO ACID	AMINO ACID POSITION#								
	1	2	3	4	5	6	7	8	9
I	8	8	8	8	8	8	8	8	8
R					8	8	8	8	8
Q	8	8	8	17	17	8	8	8	8
G						8	8	8	8
E	8								
V	8	8	8	8	8	8			
T	17	17	17	17	17	17	17	8	8
L									8
A	17	17	17	17	17	17	17	17	17
N	8	8	8	8	8	8	17	17	8
S	17	17	17	17	8	8	8	8	8
M	8	17	17	8	8	8	8	8	8
F								8	8

Figure 11

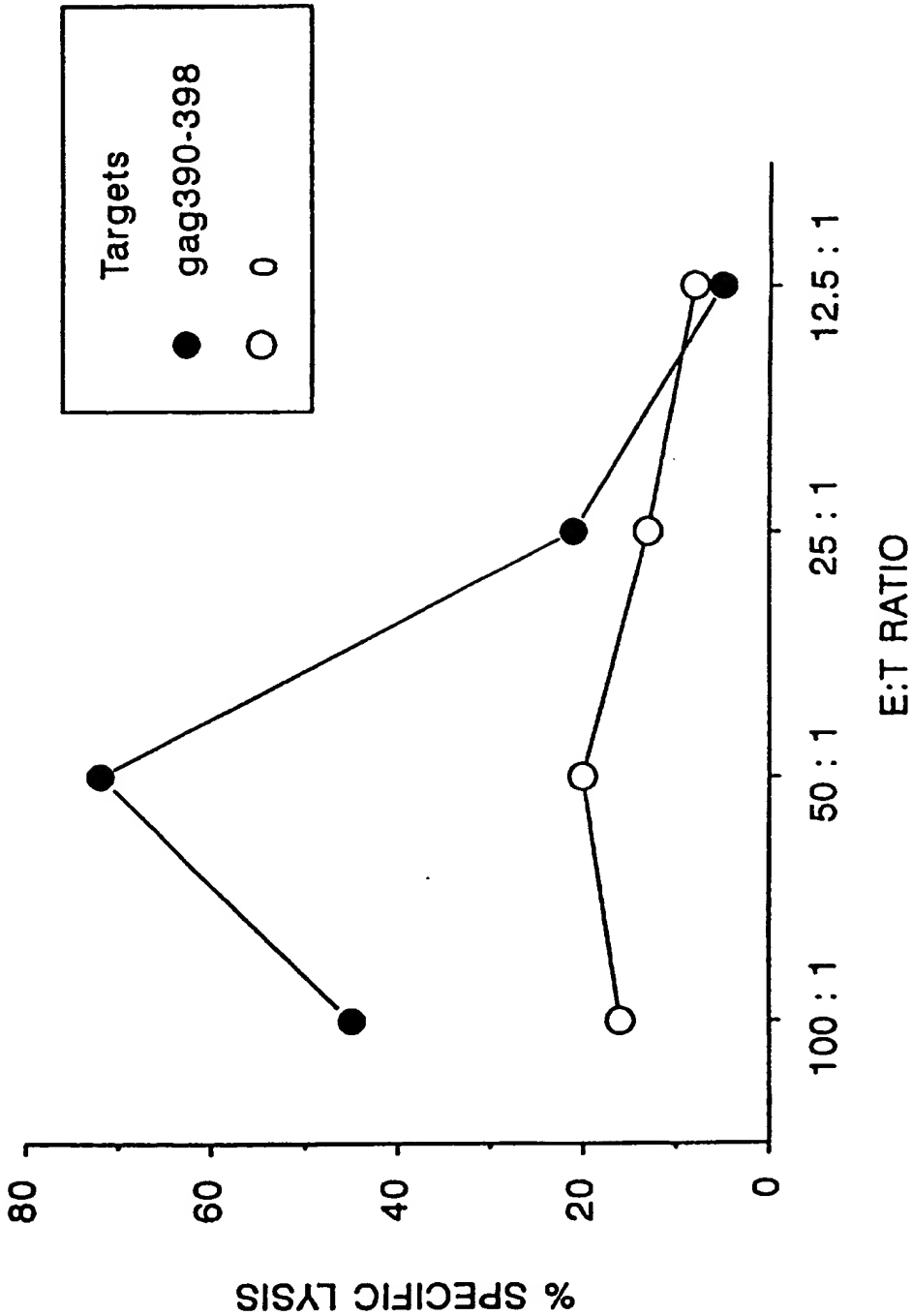


Figure 12

15/18

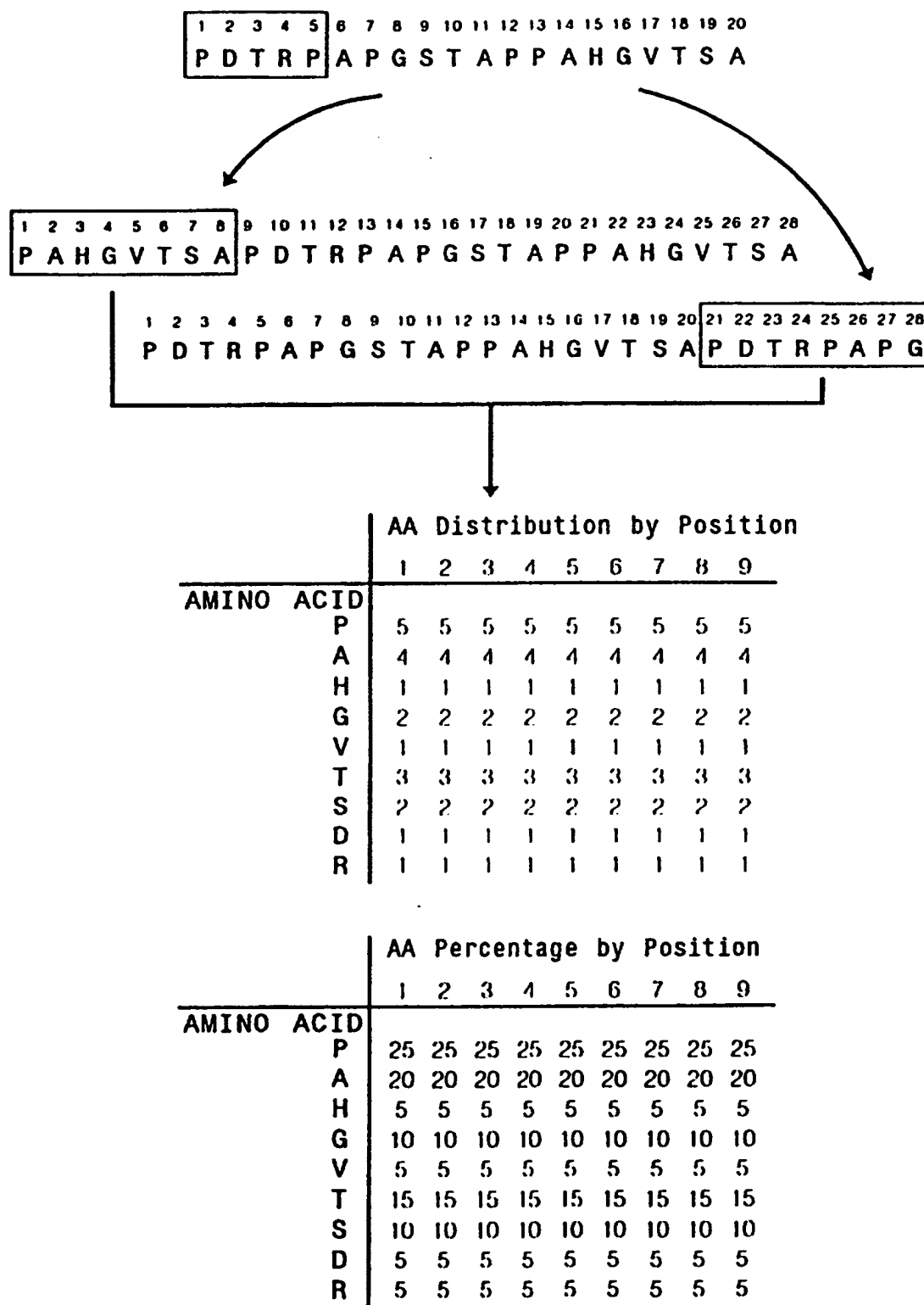


Figure 13

16/18

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28  
 C T Y S P A L N K M **F C Q L A K T C P V** Q L W V D S T P  
 M L L Y

## AA Distribution by Position

	1	2	3	4	5	6	7	8	9	10
AMINO ACID										
C	3	2	2	2	2	2	2	2	2	2
T	2	2	1	1	1	1	1	1	2	2
Y	2	2	2	1	1	1	1	1	1	1
S	1	1	1	1				1	1	1
P	2	2	2	2	2	1	1	1	1	2
A	2	2	2	2	2	2	1	1	1	1
L	4	4	4	5	5	5	5	4	4	4
N	1	1	1	1	1	1	1	1		
K	2	2	2	2	2	2	2	2	2	1
M	2	2	2	2	2	2	2	2	2	1
F	1	1	1	1	1	1	1	1	1	1
Q	1	1	2	2	2	2	2	2	2	2
V		1	1	1	1	2	2	2	2	2
W					1	1	1	1	1	1
D							1	1	1	1

## AA Percentage by Position

	1	2	3	4	5	6	7	8	9	10
AMINO ACID										
C	13	9	9	9	9	9	9	9	9	11
T	9	9	4	4	4	4	4	4	9	11
Y	9	9	9	4	4	4	4	4	4	5
S	4	4	4	4				4	4	5
P	9	9	9	9	9	4	4	4	4	11
A	9	9	9	9	9	9	4	4	4	5
L	17	17	17	22	22	22	22	17	17	18
N	4	4	4	4	4	4	4	4		
K	9	9	9	9	9	9	9	9	9	5
M	9	9	9	9	9	9	9	9	9	5
F	4	4	4	4	4	4	4	4	4	5
Q	4	4	9	9	9	9	9	9	9	11
V		4	4	4	4	9	9	9	9	11
W					4	4	4	4	4	5
D							4	4	4	5

Figure 14

17/18

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25  
W H S N L N D A **T Y Q R T R A L V** R T G M D P R M

## AA Distribution by Position

	1	2	3	4	5	6	7	8	9
AMINO ACID									
W	1								
H	1	1							
S	1	1	1						
N	2	2	2	2	1	1			
L	2	2	2	2	2	1	1	1	1
D	1	1	1	1	1	1	2	1	1
A	2	2	2	2	2	2	2	2	1
T	2	2	2	3	3	3	3	3	3
Y	1	1	1	1	1	1	1	1	1
Q	1	1	1	1	1	1	1	1	1
R	2	2	3	3	3	3	3	3	4
V		1	1	1	1	1	1	1	1
G					1	1	1	1	1
M						1	1	1	1
P								1	1

## AA Percentage by Position

	1	2	3	4	5	6	7	8	9
AMINO ACID									
W	6								
H	6	6							
S	6	6	6						
N	13	13	13	13	6	6			
L	13	13	13	13	13	6	6	6	7
D	6	6	6	6	6	6	13	6	7
A	13	13	13	13	13	13	13	13	7
T	13	13	13	19	19	19	19	19	13
Y	6	6	6	6	6	6	6	6	7
Q	6	6	6	6	6	6	6	6	7
R	13	13	19	19	19	19	19	19	25
V		6	6	6	6	6	6	6	7
G					6	6	6	6	7
M						6	6	6	7
P								6	7

Figure 15

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
S	D	Y	E	G	R	L	I	T	Y	Q	R	T	R	A	L	V	A	S	N	E	N	M	E	T	M

## AA Distribution by Position

	1	2	3	4	5	6	7	8	9
AMINO ACID									
S	1	1	1	1	1	1	1	1	1
D	2	2	1	1	1	1	1	1	1
Y	2	2	2	1	1	1	1	1	1
E	1	1	1	2	1	1	2	2	2
G	1	1	1	1	1				
R	3	3	3	3	3	3	2	2	2
L	2	2	2	2	2	2	2	1	1
I	1	1	1	1	1	1	1	1	
T	2	2	2	2	2	2	2	3	3
Q	1	1	1	1	1	1	1	1	1
A	2	2	2	2	2	2	2	2	2
V	1	1	1	1	1	1	1	1	1
N			1	1	2	2	2	2	2
M						1	1	1	2

## AA Percentage by Position

	1	2	3	4	5	6	7	8	9
AMINO ACID									
S	5	5	5	5	5	5	5	5	5
D	11	11	5	5	5	5	5	5	5
Y	11	11	11	5	5	5	5	5	5
E	5	5	5	11	5	5	11	11	11
G	5	5	5	5	5				
R	16	16	16	16	16	16	11	11	11
L	11	11	11	11	11	11	11	5	5
I	5	5	5	5	5	5	5	5	
T	11	11	11	11	11	11	11	16	16
Q	5	5	5	5	5	5	5	5	5
A	11	11	11	11	11	11	11	11	11
V	5	5	5	5	5	5	5	5	5
N			5	5	11	11	11	11	11
M						5	5	5	11

Figure 16